

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:54:57 ; Search time 44.8209 Seconds  
(without alignments)  
69.343 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
1	62	100.0	11	4	AAB72505	Aab72505 Colostrin	
2	62	100.0	11	4	AAB59311	Aab59311 Ewe colos	
3	62	100.0	11	4	AAB72251	Aab72251 Colostrin	
4	62	100.0	11	4	AAB72537	Aab72537 Colostrin	
5	62	100.0	11	5	AAO14582	Aao14582 Neural ce	
6	62	100.0	11	5	AAM51041	Aam51041 Colostrin	
7	62	100.0	11	5	AAE20233	Aae20233 Colostrin	
8	62	100.0	12	4	AAB59342	Aab59342 Ewe colos	
9	40	64.5	673	6	ABP71874	Abp71874 Growth ho	

10	40	64.5	716	4	ABG24050	Abg24050	Novel hum
11	40	64.5	716	7	ADC31679	Adc31679	Human nov
12	40	64.5	716	7	ADC33261	Adc33261	Human nov
13	39	62.9	582	2	AAy37674	Aay37674	Protein i
14	39	62.9	582	2	AAW97215	Aaw97215	An aspart
15	39	62.9	582	6	ABU27392	Abu27392	Protein e
16	39	62.9	591	5	ABB49281	Abb49281	Listeria
17	39	62.9	591	6	ABU32645	Abu32645	Protein e
18	39	62.9	3460	5	ABB05007	Abb05007	Human ree
19	39	62.9	3460	7	ADE60143	Ade60143	Human Pro
20	39	62.9	3461	5	ABB05008	Abb05008	Mouse ree
21	39	62.9	3461	5	ABB57065	Abb57065	Mouse isc
22	39	62.9	3461	7	ADE60142	Ade60142	Rat Prote
23	39	62.9	3470	4	ABG25297	Abg25297	Novel hum
24	37	59.7	114	4	ABB67449	Abb67449	Drosophil
25	37	59.7	200	4	AAU19167	Aau19167	Human G p
26	37	59.7	218	4	AAO13220	Aao13220	Human pol
27	37	59.7	433	2	AAR98454	Aar98454	Oligodend
28	37	59.7	433	3	AAy57091	Aay57091	Human oli
29	37	59.7	490	6	ADA32995	Ada32995	Acinetoba
30	37	59.7	491	2	AAW52826	Aaw52826	Human chr
31	37	59.7	834	2	AAW52820	Aaw52820	Human PRC
32	36	58.1	58	5	ABG65607	Abg65607	Human bre
33	36	58.1	62	5	ABP11495	Abp11495	Human ORF
34	36	58.1	126	2	AAR92475	Aar92475	C2GnT cat
35	36	58.1	322	6	ABU38887	Abu38887	Protein e
36	36	58.1	382	2	AAW06517	Aaw06517	Flavobact
37	36	58.1	382	2	AAW69533	Aaw69533	Flavobact
38	36	58.1	427	3	AAB30297	Aab30297	Diabetic
39	36	58.1	428	2	AAW93942	Aaw93942	Rat DH1 p
40	36	58.1	428	2	AAW93943	Aaw93943	Human cor
41	36	58.1	428	3	AAB30298	Aab30298	Human hea
42	36	58.1	428	5	AAM48976	Aam48976	Human glu
43	36	58.1	469	4	AAB96093	Aab96093	Putative
44	36	58.1	715	4	AAM78647	Aam78647	Human pro
45	36	58.1	715	4	ABU52753	Abu52753	Human nuc

# ALIGNMENTS

RESULT 1

AAB72505

ID AAB72505 standard; peptide; 11 AA.

XX

AC AAB72505;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #6.

XX

KW Dermatological; oxidative stress regulator; colostrinin.

XX

OS Unidentified.

XX

PN WO200112650-A2.

XX

PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022665.  
 XX  
 PR 17-AUG-1999; 99US-0149310P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I;  
 XX  
 DR WPI; 2001-218342/22.  
 XX  
 PT Modulating oxidative stress level in a cell, involves contacting the cell  
 PT with an oxidative stress regulator selected from colostrinin, its  
 PT constituent peptide, analog or their combinations.  
 XX  
 PS Claim 6; Page 25; 48pp; English.  
 XX  
 CC The present invention relates to a method for modulating the oxidative  
 CC stress level in a cell or a patient, comprising contacting the cell with,  
 CC or administering to the patient, an oxidative stress regulator selected  
 CC from colostrinin, or its constituent peptide (e.g. the present peptide),  
 CC to change the level of an oxidising species in the cell. The method can  
 CC be used to treat oxidative damage to skin, by decreasing or preventing an  
 CC increase in the level of damage to a biomolecule of the patient  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11

# RESULT 2

AAB59311

ID AAB59311 standard; peptide; 11 AA.

XX

AC AAB59311;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment A-2.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;  
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX  
 PR 02-JUN-1999; 99GB-00012852.  
 XX  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Georgiades JA;  
 XX  
 DR WPI; 2001-071058/08.  
 XX  
 PT Peptides having an N-terminal amino acid sequence isolated from  
 PT colostrinin for treating e.g. disorders of the central nervous system and  
 PT immune system, viral and bacterial infections, and diseases characterized  
 PT by amyloid plaques.  
 XX  
 PS Claim 7; Page 27; 63pp; English.  
 XX  
 CC The present invention provides the sequences of a number of peptides  
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide  
 CC fragment of colostrum. These peptides can be used in the treatment of  
 CC central nervous system disorders such as senile dementia, Parkinson's  
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system  
 CC disorders such as bacterial and viral infections, to improve the  
 CC development of a child's immune system, as a dietary supplement, and to  
 CC promote the dissolution of beta-amyloid plaques  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11

# RESULT 3

AAB72251

ID AAB72251 standard; peptide; 11 AA.

XX

AC AAB72251;

XX

DT 14-MAY-2001 (first entry)

XX

DE Colostrinin derived cytokine inducing peptide SEQ ID 6.

XX

KW Colostrinin; immune response; cytokine; blood cell proliferation;  
 KW central nervous system disorder; neurological disorder; mental disorder;  
 KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;  
 KW neurosis; infection.

XX

OS Synthetic.

XX

PN WO200111937-A2.

XX

PD 22-FEB-2001.

XX

PF 17-AUG-2000; 2000WO-US022818.  
 XX  
 PR 17-AUG-1999; 99US-0149311P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2001-202804/20.  
 XX  
 PT Inducing a cytokine and modulating an immune response, useful for  
 PT treating central nervous system diseases and bacterial and viral  
 PT infections, comprises administering colostrinin as an immunological  
 PT regulator.  
 XX  
 PS Claim 1; Page 34; 50pp; English.  
 XX  
 CC Sequences AAB72246 - AAB72275 represent peptides derived from clostrinin,  
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides  
 CC have immune response modulatory activity, and are capable of inducing  
 CC cytokines. Colostrinin and its derived peptides are useful for inducing  
 CC cytokine production, for modulating an immunological response and for  
 CC inducing blood cell proliferation. The peptides are useful in the  
 CC treatment of disorders of the central nervous system, neurological  
 CC disorders, mental disorders, dementia, neurodegenerative diseases,  
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic  
 CC disorders of the immune system, bacterial and viral infections and  
 CC acquired immunological deficiencies  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11

#### RESULT 4

AAB72537

ID AAB72537 standard; peptide; 11 AA.

XX

AC AAB72537;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #6.

XX

KW Neuroprotective; neural cell differentiation regulator; colostrinin;  
 KW colostrum.

XX

OS Unidentified.

XX

PN WO200112651-A2.

XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022774.  
 XX  
 PR 17-AUG-1999; 99US-0149633P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I;  
 XX  
 DR WPI; 2001-226545/23.  
 XX  
 PT Use of colostrinin, its constituent peptide or analog as a neural cell  
 PT regulator, for promoting neural cell differentiation and treating damaged  
 PT neural cells in a patient.  
 XX  
 PS Claim 6; Page 21; 35pp; English.  
 XX  
 CC The present invention relates to a method for promoting neural cell  
 CC differentiation and treating damaged neural cells, using colostrinin and  
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural  
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11

RESULT 5  
 AAO14582  
 ID AAO14582 standard; peptide; 11 AA.  
 XX  
 AC AAO14582;  
 XX  
 DT 27-MAY-2002 (first entry)  
 XX  
 DE Neural cell regulatory colostrinin peptide 6.  
 XX  
 KW Neural cell differentiation; neural cell regulator; colostrinin peptide;  
 KW neural cell formation; proline-rich polypeptide aggregate; colostrum;  
 KW neural cell treatment.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 11  
 FT /note= "Optional C-terminal amide"  
 XX  
 PN WO200213851-A1.  
 XX

PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022777.  
 XX  
 PR 17-AUG-2000; 2000WO-US022777.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I, Stanton JG, Hughes TK;  
 XX  
 DR WPI; 2002-269152/31.  
 XX  
 PT Promoting cell differentiation in a patient involves use of blood cell  
 PT regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Claim 7; Page 21; 37pp; English.  
 XX  
 CC The invention comprises a method for promoting cell differentiation (e.g.  
 CC neural cell differentiation). The method involves contacting cells with a  
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the  
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich  
 CC polypeptide aggregate that is present in colostrum. The method of the  
 CC invention is useful for promoting the differentiation of cells and for  
 CC treating damaged neural cells in a patient. The present amino acid  
 CC sequence represents a specifically claimed colostrinin peptide used in  
 CC the method of the invention  
 XX  
 SQ Sequence 11 AA;  
  
 Query Match 100.0%; Score 62; DB 5; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11  
  
 RESULT 6  
 AAM51041  
 ID AAM51041 standard; peptide; 11 AA.  
 XX  
 AC AAM51041;  
 XX  
 DT 30-MAY-2002 (first entry)  
 XX  
 DE Colostrinin constituent peptide.  
 XX  
 KW Colostrinin; colostrum; immunomodulator; cardiovascular;  
 KW blood cell regulator; cytokine inducer; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 11  
 FT /note= "optional C-terminal amidation"

XX  
 PN WO200213849-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022775.  
 XX  
 PR 17-AUG-2000; 2000WO-US022775.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2002-269150/31.  
 XX  
 PT Modulation of blood cell proliferation in a patient involves use of blood  
 PT cell regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Claim 1; Page 34; 54pp; English.  
 XX  
 CC The present sequence is that of a colostrinin constituent peptide that is  
 CC preferred for use as an immunological regulator and as a blood cell  
 CC regulator in claimed methods of the invention. Methods are claimed for:  
 CC inducing a cytokine in a cell by contact with an immunological regulator,  
 CC where the cell is present in a cell culture, a tissue, an organ or an  
 CC organism, and the cell is mammalian, including human; modulating an  
 CC immune response in a cell by contact with the immunological regulator  
 CC under conditions effective to induce a cytokine; modulating an immune  
 CC response in a patient by administering an immunological regulator under  
 CC conditions effective to induce a cytokine, where the immunological  
 CC regulator is administered topically or as part of a dietary supplement,  
 CC and where the immune response is specific or non specific, an interferon  
 CC response or an antibody response; modulating blood cell proliferation by  
 CC contacting blood cells with a blood cell regulator, where the blood cells  
 CC are present in a cell culture or an organism, are mammalian or human, and  
 CC where the blood cells are increased in number or differentiated; and a  
 CC method for modulating blood cell proliferation in a patent. A claimed  
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an  
 CC active agent such as the present peptide. Cytokines induced by this  
 CC peptide in human leucocyte cultures include interferon-gamma, tumour  
 CC necrosis factor-alpha, interleukin-6 and interleukin-10  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 5; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 |||||  
 Db 1 MPQNFYKLPQM 11

RESULT 7  
 AAE20233



ID AAE20233 standard; peptide; 11 AA.  
XX  
AC AAE20233;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Colostrinin constituent peptide #6.  
XX  
KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;  
KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;  
KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;  
KW transplantation; implantation; dermatological; vulnerary.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11  
FT /note= "Optionally C-terminal amide"  
XX  
PN WO200213850-A1.  
XX  
PD 21-FEB-2002.  
XX  
PF 17-AUG-2000; 2000WO-US022776.  
XX  
PR 17-AUG-2000; 2000WO-US022776.  
XX  
PA (TEXA ) UNIV TEXAS SYSTEM.  
XX  
PI Stanton GJ, Hughes TK, Boldogh I;  
XX  
DR WPI; 2002-269151/31.  
XX  
PT Composition useful for the modulation of blood cell proliferation in a  
PT patient comprises a blood cell regulator selected from colostrinin, its  
PT constituent peptide and/or analog.  
XX  
PS Claim 6; Page 25; 51pp; English.  
XX  
CC The invention relates to a composition which comprises a blood cell  
CC regulator selected from colostrinin, its constituent peptide and/or  
CC analogue. The invention is used for modulating the oxidative stress level  
CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,  
CC organ, or organism; or for treating oxidative damage to the skin of a  
CC patient e.g. animal or human; to modulate oxidative stress during/ after  
CC a premature birth or normal birth, preventing/delaying aging in a  
CC patient, enhancing wound healing, and the reduction of side effects of  
CC cosmetic procedures. The method changes the level of an oxidising species  
CC in the cell, such as decreases or prevents increase in the level of  
CC damage to a biomolecule of the patient selected from DNA, protein and/or  
CC lipid, compared to the same conditions when the oxidative stress  
CC regulator is not present. The modulation of oxidative stress results in  
CC enhanced repair, regeneration, and replacement of cells, tissues and  
CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and  
CC external organs), as well as enhanced preservation of such organs for  
CC transplantation, implantation, or scientific research. The present  
CC sequence is a colostrinin constituent peptide

XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 5; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
|||||||  
Db 1 MPQNFYKLPQM 11

RESULT 8

AAB59342

ID AAB59342 standard; peptide; 12 AA.

XX

AC AAB59342;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment derived sequence #2.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;  
KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX

PR 02-JUN-1999; 99GB-00012852.

XX

PA (REGE-) REGEN THERAPEUTICS PLC.

XX

PI Georgiades JA;

XX

DR WPI; 2001-071058/08.

XX

PT Peptides having an N-terminal amino acid sequence isolated from  
PT colostrinin for treating e.g. disorders of the central nervous system and  
PT immune system, viral and bacterial infections, and diseases characterized  
PT by amyloid plaques.

XX

PS Claim 8; Page 27; 63pp; English.

XX

CC The present invention provides the sequences of a number of peptides  
CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide  
CC fragment of colostrum. These peptides can be used in the treatment of  
CC central nervous system disorders such as senile dementia, Parkinson's  
CC disease, Alzheimer's disease, psychosis and neurosis, immune system  
CC disorders such as bacterial and viral infections, to improve the  
CC development of a child's immune system, as a dietary supplement, and to  
CC promote the dissolution of beta-amyloid plaques

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQNFYKLPQM 11  
| | | | | | | | | |  
Db 2 MPQNFYKLPQM 12

RESULT 9

ABP71874

ID ABP71874 standard; protein; 673 AA.

XX

AC ABP71874;

XX

DT 23-APR-2003 (first entry)

XX

DE Growth hormone 74.03.

XX

KW Growth hormone; growth hormone 74.03; cardiovascular disease;

KW development disorder; pregnancy relative disease; tumour;

KW immunological disease; inflammation.

XX

OS Unidentified.

XX

PN CN1364813-A.

XX

PD 21-AUG-2002.

XX

PF 10-JAN-2001; 2001CN-00105152.

XX

PR 10-JAN-2001; 2001CN-00105152.

XX

PA (BIOW-) BLOWINDOW GENE DEV INC SHANGHAI.

XX

PI Mao Y, Xie Y;

XX

DR WPI; 2003-000524/01.

DR N-PSDB; ABZ75217.

XX

PT New polypeptide-growth hormone 74.03 and polynucleotide for encoding such polypeptide.

PT

XX

PS Claim 1; Page 27-28 (Disclosure); 34pp; Chinese.

XX

CC The invention relates to the novel polypeptide, growth hormone 74.03, and  
CC the polynucleotide encoding it. The polypeptide is useful in treating  
CC various diseases, such as cardiovascular diseases, development disorder,  
CC pregnancy relative disease, various tumours, immunological diseases and  
CC some inflammations. The invention also discloses the antagonist resisting  
CC the polypeptide and its treatment effect, and the application of the  
CC polynucleotide. The present sequence represents the growth hormone 74.03  
CC of the invention

XX

SQ Sequence 673 AA;

Query Match 64.5%; Score 40; DB 6; Length 673;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPQNFYKLPQ 10  
:|::||| ||  
Db 353 LPKSFYKAPQ 362

RESULT 10

ABG24050

ID ABG24050 standard; protein; 716 AA.

XX

AC ABG24050;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #24041.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR N-PSDB; AAS88237.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX

PS Claim 20; SEQ ID NO 54409; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 716 AA;

Query Match 64.5%; Score 40; DB 4; Length 716;  
 Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 :|::||| ||  
 Db 406 LPKSFYKAPQ 415

# RESULT 11

ADC31679

ID ADC31679 standard; protein; 716 AA.

XX

AC ADC31679;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel polypeptide sequence, SEQ ID NO:1761.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 6q22.1-22.33.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

XX

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

DR WPI; 2003-371981/35.

DR N-PSDB; ADC30708.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.

XX

PS Claim 20; SEQ ID NO 1761; 1185pp; English.

XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human polypeptide sequence of the invention. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 716 AA;

Query Match 64.5%; Score 40; DB 7; Length 716;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
:|::||| ||  
Db 406 LPKSFYKAPQ 415

RESULT 12

ADC33261

ID ADC33261 standard; protein; 716 AA.

XX

AC ADC33261;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3343.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 6q22.1-22.33.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

XX

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

DR WPI; 2003-371981/35.

DR N-PSDB; ADC32494.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.

XX

PS Example 2; SEQ ID NO 3343; 1185pp; English.

XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

CC invention also relates to nucleic acid sequences over 99% identical with

CC the novel human cDNAs. The invention additionally encompasses expression

CC vectors and host cells comprising a nucleic acid of the invention; the

CC recombinant production of a polypeptide of the invention; an antibody

CC against a polypeptide of the invention; a method of detecting

CC polynucleotides or polypeptides of the invention; and methods of

CC identifying a compound which binds to a polypeptide of the invention. The

CC invention further discloses methods of preventing, treating or

CC ameliorating a medical condition; kits comprising polynucleotide probes

CC and/or monoclonal antibodies for carrying out the methods of the

CC invention; methods for the identification of compounds that modulate the

CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a human contig-  
CC encoded polypeptide sequence used in an example of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 716 AA;

Query Match 64.5%; Score 40; DB 7; Length 716;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
:|::||| ||  
Db 406 LPKSFYKAPQ 415

#### RESULT 13

AA37674

ID AAY37674 standard; protein; 582 AA.

XX

AC AAY37674;

XX

DT 07-OCT-1999 (first entry)

XX

DE Protein involved in intermediate metabolism of polypeptides.

XX

KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;

KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;

KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;

KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX

OS Chlamydia trachomatis.

XX

PN WO9928475-A2.

XX

PD 10-JUN-1999.

XX

PF 27-NOV-1998; 98WO-IB001939.

XX

PR 28-NOV-1997; 97FR-00015041.

PR 17-DEC-1997; 97FR-00016034.

PR 04-NOV-1998; 98US-0107077P.



XX  
 PA (GEST ) GENSET.  
 XX  
 PI Griffais R;  
 XX  
 DR WPI; 1999-371125/31.  
 XX  
 PT Genome sequence of Chlamydia trachomatis.  
 XX  
 PS Disclosure; Page 1302; 1755pp; English.  
 XX  
 CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
 CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as  
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
 CC can also be used to control growth of the microorganism. Chlamydia  
 CC trachomatis is responsible for a large number of diseases, e.g. eye  
 CC diseases such as conventional trachoma, nonendemic trachoma,  
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
 CC nongonococcal urethritis, epidymitis, cervicitis, salpingitis,  
 CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants; and  
 CC venereal lymphogranulomatosis. The polypeptides of the invention may be  
 CC of use in treating these diseases  
 XX  
 SQ Sequence 582 AA;

Query Match 62.9%; Score 39; DB 2; Length 582;  
 Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 190 PGNFYALPQ 198

RESULT 14  
 AAW97215  
 ID AAW97215 standard; protein; 582 AA.  
 XX  
 AC AAW97215;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE An aspartyl tRNA synthetase (aspS) polypeptide.  
 XX  
 KW Aspartyl tRNA synthetase; aspS; vaccine; pathogenesis; ocular trachoma;  
 KW genital trachoma; lymphogranuloma venerium; keratitis;  
 KW corneal infiltration; prostatitis; infertility; bacterial adhesion;  
 KW cell invasion.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 PN EP893494-A2.  
 XX  
 PD 27-JAN-1999.  
 XX  
 PF 10-JUL-1998; 98EP-00305501.  
 XX

PR 23-JUL-1997; 97US-00899244.

XX

PA (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX

PI Reichard R, Brown JR, Lawlor EJ;

XX

DR WPI; 1999-097779/09.

DR N-PSDB; AAX15502.

XX

PT New Chlamydia trachomatis aspartyl tRNA synthetase (aspS) polypeptide and  
PT polynucleotide - useful as diagnostic reagents and for prevention and  
PT treatment of Chlamydia trachomatis infections which cause infertility and  
PT ocular trachoma.

XX

PS Claim 6; Page 7-8; 27pp; English.

XX

CC The present sequence represents an aspartyl tRNA synthetase (aspS)  
CC polypeptide. AspS polypeptides are useful for screening for specific  
CC inhibitors or activators, in vaccine, to study their role in pathogenesis  
CC and to raise antibodies. AspS antagonists are also useful for treating  
CC conditions requiring inhibition of aspS, specifically infection by C.  
CC trachomatis which cause ocular or genital trachoma, lymphogranuloma  
CC venerium, keratitis, corneal infiltration, prostatitis and infertility  
CC etc. AspS polypeptides, polynucleotides and antagonists can also be used  
CC to prevent adhesion of bacteria, especially Gram-positive, to  
CC extracellular matrix proteins in wounds or on in-dwelling devices, and to  
CC block aspS polypeptide-mediated cell invasion or normal progression of  
CC infection

XX

SQ Sequence 582 AA;

Query Match 62.9%; Score 39; DB 2; Length 582;

Best Local Similarity 77.8%; Pred. No. 1.7e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10

| ||| |||

Db 190 PGNFYALPQ 198

RESULT 15

ABU27392

ID ABU27392 standard; protein; 582 AA.

XX

AC ABU27392;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #12919.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Chlamydia trachomatis.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI; 2003-029926/02.  
DR N-PSDB; ACA31262.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 55316; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 582 AA;

Query Match 62.9%; Score 39; DB 6; Length 582;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| ||| |||  
Db 190 PGNFYALPQ 198

Search completed: August 24, 2004, 15:42:31  
Job time : 47.8209 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:33:13 ; Search time 12.0672 Seconds  
(without alignments)  
47.060 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	11	4	US-09-641-803-6	Sequence 6, Appli
2	39	62.9	582	2	US-08-899-244-2	Sequence 2, Appli
3	39	62.9	582	3	US-09-224-772-2	Sequence 2, Appli
4	39	62.9	582	3	US-09-227-804-2	Sequence 2, Appli
5	39	62.9	3460	4	US-09-334-220-1	Sequence 1, Appli
6	39	62.9	3461	4	US-09-334-220-2	Sequence 2, Appli
7	37	59.7	490	4	US-09-328-352-4282	Sequence 4282, Ap
8	36	58.1	126	1	US-08-118-906-4	Sequence 4, Appli
9	36	58.1	126	1	US-08-486-196-4	Sequence 4, Appli
10	36	58.1	126	1	US-08-488-135-4	Sequence 4, Appli
11	36	58.1	126	2	US-08-474-065-4	Sequence 4, Appli

12	36	58.1	211	4	US-09-134-000C-5667	Sequence 5667, Ap
13	36	58.1	360	4	US-09-252-991A-23023	Sequence 23023, A
14	36	58.1	382	3	US-08-660-645A-7	Sequence 7, Appli
15	36	58.1	382	3	US-09-298-718-7	Sequence 7, Appli
16	36	58.1	382	3	US-09-546-969-7	Sequence 7, Appli
17	36	58.1	382	3	US-08-980-832-5	Sequence 5, Appli
18	36	58.1	382	4	US-09-547-267-7	Sequence 7, Appli
19	36	58.1	382	4	US-09-920-923B-5	Sequence 5, Appli
20	36	58.1	428	1	US-07-955-041-4	Sequence 4, Appli
21	36	58.1	428	1	US-08-227-455-4	Sequence 4, Appli
22	36	58.1	428	1	US-08-472-482-4	Sequence 4, Appli
23	36	58.1	428	1	US-08-487-069-4	Sequence 4, Appli
24	36	58.1	428	3	US-09-233-506-3	Sequence 3, Appli
25	36	58.1	428	4	US-09-645-192-13	Sequence 13, Appl
26	36	58.1	1038	4	US-09-081-385-151	Sequence 151, App
27	35	56.5	45	4	US-09-257-179-64	Sequence 64, Appl
28	35	56.5	464	4	US-09-543-681A-5655	Sequence 5655, Ap
29	35	56.5	542	4	US-09-589-733C-2	Sequence 2, Appli
30	35	56.5	542	4	US-09-589-733C-20	Sequence 20, Appl
31	35	56.5	761	4	US-09-625-188-14	Sequence 14, Appl
32	35	56.5	1042	3	US-09-387-695-2	Sequence 2, Appli
33	34.5	55.6	596	3	US-08-481-190-8	Sequence 8, Appli
34	34.5	55.6	596	5	PCT-US93-00869-8	Sequence 8, Appli
35	34	54.8	235	1	US-07-971-160-4	Sequence 4, Appli
36	34	54.8	235	1	US-08-336-241-4	Sequence 4, Appli
37	34	54.8	235	2	US-08-465-273-4	Sequence 4, Appli
38	34	54.8	235	2	US-09-119-024-4	Sequence 4, Appli
39	34	54.8	235	2	US-08-417-226-4	Sequence 4, Appli
40	34	54.8	235	3	US-09-196-131-4	Sequence 4, Appli
41	34	54.8	235	4	US-08-643-732-4	Sequence 4, Appli
42	34	54.8	235	4	US-09-836-169-4	Sequence 4, Appli
43	34	54.8	237	1	US-07-971-160-2	Sequence 2, Appli
44	34	54.8	237	1	US-07-971-160-16	Sequence 16, Appl
45	34	54.8	237	1	US-08-336-241-2	Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-641-803-6

; Sequence 6, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641,803

; CURRENT FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-09-641-803-6

Query Match 100.0%; Score 62; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
| | | | | | | | | |  
Db 1 MPQNFYKLPQM 11

RESULT 2

US-08-899-244-2

; Sequence 2, Application US/08899244  
; Patent No. 5882892  
; GENERAL INFORMATION:  
; APPLICANT: Reichard, Raymond W.  
; APPLICANT: Brown, James R.  
; APPLICANT: Lawlor, Elizabeth J.  
; TITLE OF INVENTION: NOVEL aspS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/899,244  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: GM10049  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 582 amino acids

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-899-244-2

Query Match 62.9%; Score 39; DB 2; Length 582;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| | | | |  
Db 190 PGNFYALPQ 198

RESULT 3

US-09-224-772-2

; Sequence 2, Application US/09224772  
; Patent No. 6207162  
; GENERAL INFORMATION:  
; APPLICANT: Reichard, Raymond W.  
; APPLICANT: Brown, James R.  
; APPLICANT: Lawlor, Elizabeth J.  
; TITLE OF INVENTION: NOVEL aspS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/224,772  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/899,244  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: GM10049  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 582 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear



US-09-224-772-2

Query Match 62.9%; Score 39; DB 3; Length 582;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| ||| |||  
Db 190 PGNFYALPQ 198

RESULT 4

US-09-227-804-2

; Sequence 2, Application US/09227804  
; Patent No. 6265188  
; GENERAL INFORMATION:  
; APPLICANT: Reichard, Raymond W.  
; APPLICANT: Brown, James R.  
; APPLICANT: Lawlor, Elizabeth J.  
; TITLE OF INVENTION: NOVEL aspS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/227,804  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/899,244  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: GM10049  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 582 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-227-804-2

Query Match 62.9%; Score 39; DB 3; Length 582;

Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| ||| |||  
Db 190 PGNFYALPQ 198

RESULT 5

US-09-334-220-1

; Sequence 1, Application US/09334220  
; Patent No. 6323177  
; GENERAL INFORMATION:  
; APPLICANT: St. Judes Children's Research Hospital  
; APPLICANT: Curran, Thomas  
; APPLICANT: D'Arcangelo, Gabriella  
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW  
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND  
; TITLE OF INVENTION: THERAPIES  
; FILE REFERENCE: 2427/0F704  
; CURRENT APPLICATION NUMBER: US/09/334,220  
; CURRENT FILING DATE: 1999-06-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 3460  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-334-220-1

Query Match 62.9%; Score 39; DB 4; Length 3460;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
:|||||: |  
Db 1235 LPQNFYEKP 1243

RESULT 6

US-09-334-220-2

; Sequence 2, Application US/09334220  
; Patent No. 6323177  
; GENERAL INFORMATION:  
; APPLICANT: St. Judes Children's Research Hospital  
; APPLICANT: Curran, Thomas  
; APPLICANT: D'Arcangelo, Gabriella  
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW  
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND  
; TITLE OF INVENTION: THERAPIES  
; FILE REFERENCE: 2427/0F704  
; CURRENT APPLICATION NUMBER: US/09/334,220  
; CURRENT FILING DATE: 1999-06-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 3461

; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-334-220-2

Query Match 62.9%; Score 39; DB 4; Length 3461;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
:|||||: |  
Db 1236 LPQNFYEKP 1244

RESULT 7

US-09-328-352-4282  
; Sequence 4282, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4282  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4282

Query Match 59.7%; Score 37; DB 4; Length 490;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
|| :||||  
Db 140 PQTLHRLPQM 149

RESULT 8

US-08-118-906-4  
; Sequence 4, Application US/08118906  
; Patent No. 5484590  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Minoru  
; APPLICANT: Bierhuizen, Marti F.A.  
; TITLE OF INVENTION: Expression of the Developmental I  
; TITLE OF INVENTION: Antigen By a Cloned Human cDNA Encoding a Member of a  
; TITLE OF INVENTION: Beta-1,6-N-Acetylglucosaminyltransferase Gene Family  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California

```

; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,906
; FILING DATE: 09-SEP-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-118-906-4

```

```

Query Match          58.1%; Score 36; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 22;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 MPQNFY 6
        |||||
Db      37 MPQNFY 42

```

# RESULT 9

US-08-486-196-4

```

; Sequence 4, Application US/08486196
; Patent No. 5731420
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Minoru
; APPLICANT: Bierhuizen, Marti F.A.
; TITLE OF INVENTION: Expression of the Developmental I
; TITLE OF INVENTION: Antigen By a Cloned Human cDNA Encoding a Member of a
; TITLE OF INVENTION: Beta-1,6-N-Acetylglucosaminyltransferase Gene Family
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/486,196
;   FILING DATE:
;   CLASSIFICATION:  424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 08/118,906
;   FILING DATE:  09-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Campbell, Cathryn A.
;   REGISTRATION NUMBER:  31,815
;   REFERENCE/DOCKET NUMBER:  P-LJ 9526
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (619) 535-9001
;   TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  126 amino acids
;   TYPE:  amino acid
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-486-196-4

```

```

Query Match          58.1%;  Score 36;  DB 1;  Length 126;
Best Local Similarity 100.0%;  Pred. No. 22;
Matches      6;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 MPQNFY 6
        |||||
Db      37 MPQNFY 42

```

# RESULT 10

US-08-488-135-4

```

; Sequence 4, Application US/08488135
; Patent No. 5766910
; GENERAL INFORMATION:
;   APPLICANT:  Fukuda, Minoru
;   APPLICANT:  Bierhuizen, Marti F.A.
;   TITLE OF INVENTION:  Expression of the Developmental I
;   TITLE OF INVENTION:  Antigen By a Cloned Human cDNA Encoding a Member of a
;   TITLE OF INVENTION:  Beta-1,6-N-Acetylglucosaminyltransferase Gene Family
;   NUMBER OF SEQUENCES:  14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Campbell and Flores
;   STREET:  4370 La Jolla Village Drive, Suite 700
;   CITY:  San Diego
;   STATE:  California
;   COUNTRY:  USA
;   ZIP:  92122
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/488,135

```

```

; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/118,906
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-135-4

```

```

Query Match          58.1%; Score 36; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 22;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 MPQNFY 6
        |||||
Db      37 MPQNFY 42

```

# RESULT 11

```

US-08-474-065-4
; Sequence 4, Application US/08474065
; Patent No. 5830465
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Minoru
; APPLICANT: Bierhuizen, Marti F.A.
; TITLE OF INVENTION: Expression of the Developmental I
; TITLE OF INVENTION: Antigen By a Cloned Human cDNA Encoding a Member of a
; TITLE OF INVENTION: Beta-1,6-N-Acetylglucosaminyltransferase Gene Family
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,065
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/118,906
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-065-4

```

```

Query Match          58.1%; Score 36; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 22;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MPQNFY 6
        |||||
Db      37 MPQNFY 42

```

# RESULT 12

US-09-134-000C-5667

```

; Sequence 5667, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5667
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5667

```

```

Query Match          58.1%; Score 36; DB 4; Length 211;
Best Local Similarity 66.7%; Pred. No. 38;
Matches      6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MPQNFYKLP 9
        :|:| |||
Db      202 IPENDYKLP 210

```

# RESULT 13

US-09-252-991A-23023  
; Sequence 23023, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23023  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23023

Query Match 58.1%; Score 36; DB 4; Length 360;  
Best Local Similarity 62.5%; Pred. No. 66;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQNFYKLP 9  
|::||:|  
Db 270 PEDFYRLP 277

#### RESULT 14

US-08-660-645A-7  
; Sequence 7, Application US/08660645A  
; Patent No. 6087152  
; GENERAL INFORMATION:  
; APPLICANT: Hohmann, Hans-Peter  
; APPLICANT: Pasamontes, Luis  
; APPLICANT: Tessier, Michel  
; APPLICANT: van Loon, Adolphus  
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/660,645A  
; FILING DATE: 07-JUN-1996



```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95108888.9
; FILING DATE: 09-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Pokras, Bruce A.
; REGISTRATION NUMBER: 32,748
; REFERENCE/DOCKET NUMBER: RAN 6002/170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-5801
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-660-645A-7

```

```

Query Match          58.1%; Score 36; DB 3; Length 382;
Best Local Similarity 75.0%; Pred. No. 70;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 QNFYKLPQ 10
        | ||:|
Db      330 QRFYRLPQ 337

```

# RESULT 15

US-09-298-718-7

```

; Sequence 7, Application US/09298718
; Patent No. 6124113
; GENERAL INFORMATION:
; APPLICANT: Hohmann, Hans-Peter
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tessier, Michel
; APPLICANT: van Loon, Adolphus
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/298,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/660,645  
;  
; FILING DATE:  
;  
; ATTORNEY/AGENT INFORMATION:  
;  
; NAME: Pokras, Bruce A.  
;  
; REGISTRATION NUMBER: 32,748  
;  
; REFERENCE/DOCKET NUMBER: RAN 6002/170  
;  
; TELECOMMUNICATION INFORMATION:  
;  
; TELEPHONE: (201) 235-5801  
;  
; TELEFAX: (201) 235-2363  
;  
; INFORMATION FOR SEQ ID NO: 7:  
;  
; SEQUENCE CHARACTERISTICS:  
;  
; LENGTH: 382 amino acids  
;  
; TYPE: amino acid  
;  
; STRANDEDNESS: single  
;  
; TOPOLOGY: linear  
;  
; MOLECULE TYPE: protein  
US-09-298-718-7

Query Match 58.1%; Score 36; DB 3; Length 382;  
Best Local Similarity 75.0%; Pred. No. 70;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QNFYKLPQ 10  
| ||:||||  
Db 330 QRFYRLPQ 337

Search completed: August 24, 2004, 15:55:17  
Job time : 14.0672 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:26:28 ; Search time 10.6716 Seconds  
(without alignments)  
99.151 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	41	66.1	113	2	B72629	hypothetical prote
2	39	62.9	247	2	H70432	ABC transporter -
3	39	62.9	582	2	G71500	aspartate-tRNA lig
4	39	62.9	582	2	E81660	aspartyl-tRNA synt
5	39	62.9	591	2	AG1264	aspartyl-tRNA synt
6	39	62.9	591	2	AI1626	aspartyl-tRNA synt
7	39	62.9	3461	2	S58870	reelin precursor -
8	38	61.3	289	2	C96752	probable DNA-bindi
9	38	61.3	445	2	B71642	hypothetical prote
10	38	61.3	445	2	F97855	hypothetical prote
11	38	61.3	1374	2	S62524	probable RNA helic
12	37	59.7	241	2	JH0365	Balbani ring a 67
13	37	59.7	266	2	S75992	hypothetical prote

14	37	59.7	440	2	F84955	UDP-N-acetylmuramo
15	37	59.7	440	2	A39613	oligodendrocyte-my
16	37	59.7	609	2	G87483	aspartyl-tRNA synt
17	37	59.7	660	1	G6BPT4	baseplate protein
18	37	59.7	779	2	T21021	hypothetical prote
19	36	58.1	184	2	S26046	sex-determining pr
20	36	58.1	322	2	C82991	probable aromatic
21	36	58.1	428	2	A46293	beta-1,3-galactosy
22	36	58.1	478	2	T23792	hypothetical prote
23	36	58.1	486	2	T21481	hypothetical prote
24	36	58.1	564	2	A86145	hypothetical prote
25	36	58.1	606	2	E87318	hypothetical prote
26	36	58.1	1186	2	T03180	tyrosine protein k
27	36	58.1	1953	2	S63244	BN11 protein - yea
28	35	56.5	182	2	S51802	vomeronal secret
29	35	56.5	195	2	B70151	endonuclease precu
30	35	56.5	210	2	T04316	heat shock protein
31	35	56.5	253	2	E83107	conserved hypothet
32	35	56.5	349	2	AH2858	glycosyltransferas
33	35	56.5	363	2	F97635	alpha-D-mannose-al
34	35	56.5	442	2	T02620	hypothetical prote
35	35	56.5	566	2	T46219	hypothetical prote
36	35	56.5	693	2	AF2275	cellulose synthase
37	35	56.5	1042	2	E95103	DNA polymerase III
38	35	56.5	1042	2	C97971	DNA-directed DNA p
39	35	56.5	1070	2	T34385	hypothetical prote
40	35	56.5	1072	2	T50949	verprolin related
41	34.5	55.6	527	2	S46088	hypothetical prote
42	34.5	55.6	527	2	S64060	probable membrane
43	34.5	55.6	596	1	S33540	catechol oxidase (
44	34.5	55.6	599	2	T07097	catechol oxidase (
45	34	54.8	62	2	AI2652	hypothetical prote

## ALIGNMENTS

### RESULT 1

B72629

hypothetical protein APE1490 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000

C;Accession: B72629

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba, S.; Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai, Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura, Y.; Nomura, N.; Sako, Y.; Kikuchi, H.

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix K1.

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: B72629

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-113 <KAW>

A;Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80488.1; PID:d1044274;  
PID:g5105174  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE1490  
C;Superfamily: Aeropyrum pernix hypothetical protein APE1490

Query Match 66.1%; Score 41; DB 2; Length 113;  
Best Local Similarity 54.5%; Pred. No. 1.4;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
:| ||: ||:|  
Db 15 LPSNFFSLPRM 25

RESULT 2  
H70432

ABC transporter - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 02-Feb-2001

C;Accession: H70432

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.; Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short, J.M.; Olson, G.J.; Swanson, R.V.

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: H70432

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-247 <AQF>

A;Cross-references: GB:AE000744; NID:g2983891; PIDN:AAC07448.1; PID:g2983897;  
GB:AE000657

A;Experimental source: strain VF5

C;Genetics:

A;Gene: abcT10

C;Superfamily: inner membrane protein malK; ATP-binding cassette homology

C;Keywords: ATP; nucleotide binding; P-loop

F;18-215/Domain: ATP-binding cassette homology <ABC>

F;35-42/Region: nucleotide-binding motif A (P-loop)

Query Match 62.9%; Score 39; DB 2; Length 247;  
Best Local Similarity 66.7%; Pred. No. 7.7;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
|::||| ||  
Db 130 PEDFYKYPQ 138

RESULT 3  
G71500

aspartate-tRNA ligase (EC 6.1.1.12) - Chlamydia trachomatis (serotype D, strain UW3/Cx)

N;Alternate names: aspartyl-tRNA synthetase

C;Species: *Chlamydia trachomatis*  
 C;Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 03-Jun-2002  
 C;Accession: G71500  
 R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W. Science 282, 754-759, 1998  
 A;Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*.  
 A;Reference number: A71570; MUID:99000809; PMID:9784136  
 A;Accession: G71500  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-582 <ARN>  
 A;Cross-references: GB:AE001325; GB:AE001273; NID:g3328980; PIDN:AAC68144.1; PID:g3328981  
 A;Experimental source: serotype D, strain UW-3/Cx  
 C;Genetics:  
 A;Gene: aspS  
 C;Function:  
 A;Description: activates amino acid and transfers it to specific tRNA molecule  
 A;Pathway: protein biosynthesis  
 C;Superfamily: lysine-tRNA ligase  
 C;Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 62.9%; Score 39; DB 2; Length 582;  
 Best Local Similarity 77.8%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 190 PGNFYALPQ 198

#### RESULT 4

E81660

aspartyl-tRNA synthetase TC0829 [imported] - *Chlamydia muridarum* (strain Nigg)

C;Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-May-2000

C;Accession: E81660

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, E.K.; Peterson, J.; Utterback, T.; Berry, K.; Bass, S.; Linher, K.; Weidman, J.; Khouri, H.; Craven, B.; Bowman, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; Eisen, J.; Fraser, C.M.

Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.

A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: E81660

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-582 <TET>

A;Cross-references: GB:AE002349; GB:AE002160; NID:g7190851; PIDN:AAF39629.1; PID:g7190856; GSPDB:GN00121; TIGR:TC0829

A;Experimental source: strain Nigg (MoPn)

C;Genetics:

A;Gene: TC0829

C;Superfamily: lysine-tRNA ligase

Query Match 62.9%; Score 39; DB 2; Length 582;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| | | | |  
Db 190 PGNFYALPQ 198

RESULT 5

AG1264

aspartyl-tRNA synthetase [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AG1264

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AG1264

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-591 <GLA>

A;Cross-references: GB:NC\_003210; PIDN:CAC99597.1; PID:g16410948; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: aspS

C;Superfamily: lysine-tRNA ligase

Query Match 62.9%; Score 39; DB 2; Length 591;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| | | | |  
Db 189 PGNFYALPQ 197

RESULT 6

AI1626

aspartyl-tRNA synthetase [imported] - *Listeria innocua* (strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

C;Accession: AI1626

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,

E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.  
 A;Title: Comparative genomics of *Listeria* species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AI1626  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-591 <GLA>  
 A;Cross-references: GB:AL592022; PIDN:CAC96785.1; PID:g16414041; GSPDB:GN00178  
 A;Experimental source: strain Clip11262  
 C;Genetics:  
 A;Gene: aspS  
 C;Superfamily: lysine-tRNA ligase

Query Match 62.9%; Score 39; DB 2; Length 591;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | | | | |  
 Db 189 PGNFYALPQ 197

# RESULT 7

S58870

reelin precursor - mouse

C;Species: *Mus musculus* (house mouse)

C;Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 21-Jul-2003

C;Accession: S58870; S71844; I49297

R;D'Arcangelo, G.; Miao, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.I.; Curran, T.

Nature 374, 719-723, 1995

A;Title: A protein related to extracellular matrix proteins deleted in the mouse mutant reeler.

A;Reference number: I49297; MUID:95231649; PMID:7715726

A;Accession: S58870

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-3461 <DAR>

A;Cross-references: EMBL:U24703; NID:g902486; PID:g902487

R;D'Arcangelo, G.

submitted to the EMBL Data Library, April 1995

A;Reference number: S71844

A;Accession: S71844

A;Molecule type: mRNA

A;Residues: 1-215, 'T', 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>

A;Cross-references: EMBL:U24703; NID:g902486; PID:g902487

F;1-27/Domain: signal sequence #status predicted <SIG>

F;28-3461/Product: reelin #status predicted <MAT>

F;1769-1795/Domain: EGF homology <EGF>



Query Match 62.9%; Score 39; DB 2; Length 3461;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
:||||: |  
Db 1236 LPQNFYEKP 1244

RESULT 8

C96752

probable DNA-binding protein F28P22.7 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C;Accession: C96752

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;  
Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.;  
Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.;  
Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.;  
Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.;  
Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.;  
Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-  
Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu,  
S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda,  
M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.;  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;  
Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken,  
S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.;  
Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C96752

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-289 <STO>

A;Cross-references: GB:AE005173; NID:g6648154; PIDN:AAF21154.1; GSPDB:GN00141

C;Genetics:

A;Gene: F28P22.7

A;Map position: 1

Query Match 61.3%; Score 38; DB 2; Length 289;  
Best Local Similarity 85.7%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLP 9  
||||: |  
Db 186 QNFYKIP 192

RESULT 9

B71642

hypothetical protein RP808 - Rickettsia prowazekii

C;Species: Rickettsia prowazekii

C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C;Accession: B71642  
 R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.;  
 Alsmark, U.C.M.; Podowski, R.M.; Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.;  
 Kurland, C.G.  
 Nature 396, 133-140, 1998  
 A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of  
 mitochondria.  
 A;Reference number: A71630; MUID:99039499; PMID:9823893  
 A;Accession: B71642  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-445 <AND>  
 A;Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15234.1;  
 PID:g3861335; GSPDB:GN00081  
 A;Experimental source: strain Madrid E  
 C;Genetics:  
 A;Gene: RP808  
 C;Superfamily: conserved hypothetical protein b0835

Query Match 61.3%; Score 38; DB 2; Length 445;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
 ||::| ||::  
 Db 109 PQSYYNLPEL 118

# RESULT 10

F97855

hypothetical protein RC1246 [imported] - *Rickettsia conorii* (strain Malish 7)

C;Species: *Rickettsia conorii*

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C;Accession: F97855

R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.;  
 Samson, D.; Roux, V.; Cossart, P.; Weissenbach, J.; Claverie, J.M.; Raoult, D.  
 Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia*  
*prowazekii*.

A;Reference number: A97700; MUID:21442074; PMID:11557893

A;Accession: F97855

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-445 <KUR>

A;Cross-references: GB:AE006914; PIDN:AAL03784.1; PID:g15620380; GSPDB:GN00173

C;Genetics:

A;Gene: RC1246

C;Superfamily: conserved hypothetical protein b0835

Query Match 61.3%; Score 38; DB 2; Length 445;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
 ||::| ||::  
 Db 109 PQSYYNLPEL 118

RESULT 11

S62524

probable RNA helicase/ribonuclease SPAC8A4.08c - fission yeast

(Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Nov-2000

C;Accession: T39130; T41432; T41192; S62524

R;Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, November 1995

A;Reference number: Z21830

A;Accession: T39130

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1374 <LYE>

A;Cross-references: EMBL:Z66569; NID:g1052533; PIDN:CAA91518.1; PID:g1052539;

SPDB:SPAC8A4.08c

R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z21993

A;Accession: T41432

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1374 <SEE>

A;Cross-references: EMBL:AL032824; PIDN:CAB37423.1; GSPDB:GN00068;

SPDB:SPCC584.10c

A;Experimental source: strain 972h-; cosmid c584

R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.

submitted to the EMBL Data Library, April 1999

A;Reference number: Z21976

A;Accession: T41192

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 722-1374 <LYN>

A;Cross-references: EMBL:AL049662; PIDN:CAB41233.1; GSPDB:GN00068;

SPDB:SPCC188.13c

C;Genetics: <LYE1>

A;Gene: SPAC8A4.08c

A;Map position: 1

C;Genetics: <SEE1>

A;Gene: SPCC584.10c

A;Map position: 3

F;32-39/Region: nucleotide-binding motif A (P-loop) #status atypical

F;141-146/Region: nucleotide-binding motif B

F;145-148/Region: DEAD/H motif #status atypical

Query Match 61.3%; Score 38; DB 2; Length 1374;

Best Local Similarity 75.0%; Pred. No. 75;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLPQ 10

|||:|||:

Db 779 QNFFKLPE 786

RESULT 12

JH0365

Balbiani ring a 67K protein - midge (*Chironomus thummi*)

C;Species: *Chironomus thummi*

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 21-Jul-2000

C;Accession: JH0365

R;Bogachev, S.S.; Blinov, A.G.; Kolesnikov, N.N.; Scherbik, S.V.; Taranin, A.V.; Sebeleva, T.E.; Baiborodin, S.I.; Kiknadze, I.I.

Gene 96, 241-247, 1990

A;Title: A tissue-specific puff (balbiani ring a) in *Chironomus thummi* may contain a gene encoding a 67kDa protein which exhibits non-tissue-specific expression.

A;Reference number: JH0365; MUID:91099682; PMID:2269433

A;Accession: JH0365

A;Molecule type: DNA

A;Residues: 1-241 <BOG>

A;Cross-references: GB:M63152; NID:g156521; PIDN:AAA62726.1; PID:g156522

A;Note: the authors translated the codon GAT for residue 216 as Thr

C;Superfamily: unassigned Balbiani ring proteins

Query Match	59.7%;	Score 37;	DB 2;	Length 241;
Best Local Similarity	75.0%;	Pred. No. 18;		
Matches	6;	Conservative	2;	Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLPQ 10

:||:||||

Db 65 KNFFKLPQ 72

RESULT 13

S75992

hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C;Species: *Synechocystis* sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 02-Feb-2001

C;Accession: S75992

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S75992

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-266 <KAN>

A;Cross-references: EMBL:D64006; GB:AB001339; NID:g1001291; PIDN:BAA10839.1; PID:d1011490; PID:g1001352

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

C;Keywords: ATP; nucleotide binding; P-loop

F;47-242/Domain: ATP-binding cassette homology <ABC>

F;64-71/Region: nucleotide-binding motif A (P-loop)

Query Match 59.7%; Score 37; DB 2; Length 266;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QNFYKLPQM 11  
| || ||||  
Db 116 QQFYLLPQM 124

RESULT 14

F84955

UDP-N-acetylmuramoylalanine-D-glutamate ligase (EC 6.3.2.9) [imported] -  
Buchnera sp. (strain APS)

C;Species: Buchnera sp.

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 03-Jun-2002

C;Accession: F84955

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A;Title: Genome sequence of the endocellular bacterial symbiont of aphids  
Buchnera sp. APS.

A;Reference number: A84930; MUID:20445173; PMID:10993077

A;Accession: F84955

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-440 <STO>

A;Cross-references: GB:AP000398; GSPDB:GN00144

A;Experimental source: strain APS

C;Genetics:

A;Gene: murD; BU218

C;Superfamily: UDP-N-acetylmuramate-alanine ligase

C;Keywords: ligase

Query Match 59.7%; Score 37; DB 2; Length 440;  
Best Local Similarity 66.7%; Pred. No. 34;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| || |:||  
Db 42 PSNFIKIPQ 50

RESULT 15

A39613

oligodendrocyte-myelin glycoprotein precursor - human

C;Species: Homo sapiens (man)

C;Date: 13-Sep-1991 #sequence\_revision 03-Oct-1995 #text\_change 31-Jan-2000

C;Accession: A36688; A34210; A39613; A30187

R;Mikol, D.D.; Alexakos, M.J.; Bayley, C.A.; Lemons, R.S.; Le Beau, M.M.;  
Stefansson, K.

J. Cell Biol. 111, 2673-2679, 1990

A;Title: Structure and chromosomal localization of the gene for the  
oligodendrocyte-myelin glycoprotein.

A;Reference number: A36688; MUID:91115958; PMID:2277079

A;Accession: A36688

A;Molecule type: DNA

A;Residues: 1-440 <MI1>

A;Cross-references: GB:X57436; NID:g23105; PIDN:CAA40684.1; PID:g23106  
R;Mikol, D.D.; Gulcher, J.R.; Stefansson, K.  
J. Cell Biol. 110, 471-479, 1990  
A;Title: The oligodendrocyte-myelin glycoprotein belongs to a distinct family of proteins and contains the HNK-1 carbohydrate.  
A;Reference number: A34210; MUID:90130636; PMID:1688857  
A;Accession: A34210  
A;Molecule type: DNA  
A;Residues: 8-20,'X',22-440 <MI2>  
A;Cross-references: GB:X51694; NID:g35144; PIDN:CAA35991.1; PID:g35145  
R;Viskochil, D.; Cawthon, R.; O'Connell, P.; Xu, G.; Stevens, J.; Culver, M.; Carey, J.; White, R.  
Mol. Cell. Biol. 11, 906-912, 1991  
A;Title: The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within the neurofibromatosis type 1 gene.  
A;Reference number: A39613; MUID:91117257; PMID:1899288  
A;Accession: A39613  
A;Molecule type: mRNA  
A;Residues: 1-20,'X',22-440 <VIS>  
A;Cross-references: GB:M63623; NID:g189385; PIDN:AAA59970.1; PID:g189386  
R;Mikol, D.D.; Stefansson, K.  
J. Cell Biol. 106, 1273-1279, 1988  
A;Title: A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in central nervous system myelin and on oligodendrocytes.  
A;Reference number: A30187; MUID:88198371; PMID:3283151  
A;Accession: A30187  
A;Molecule type: protein  
A;Residues: 25-29,'H',31,'H',33-44,'D',46-50,'P',52 <MI3>  
C;Genetics:  
A;Gene: GDB:OMG; OMGP  
A;Cross-references: GDB:127563; OMIM:164345  
A;Map position: 17q11.2-17q11.2  
A;Introns: #status absent  
A;Note: this gene lies within an intron of GDB:NF1 on the opposite strand  
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology  
C;Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkage; phosphoprotein; tandem repeat  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;56-78/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F;79-99/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F;100-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F;124-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F;147-167/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F;168-191/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F;192-215/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F;216-227/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F;426-440/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F;425/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in mature form) #status predicted

Query Match 59.7%; Score 37; DB 2; Length 440;  
Best Local Similarity 60.0%; Pred. No. 34;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPQNFYKLPQ 10  
|| || ::||  
Db 386 MPNNFSEMPQ 395

Search completed: August 24, 2004, 15:52:54  
Job time : 13.6716 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:51:19 ; Search time 39.8134 Seconds  
(without alignments)  
86.825 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description



1	62	100.0	11	14	US-10-281-652-6	Sequence 6, Appli
2	40	64.5	164	14	US-10-029-386-32258	Sequence 32258, A
3	40	64.5	511	15	US-10-108-260A-3071	Sequence 3071, Ap
4	39	62.9	582	12	US-10-282-122A-55316	Sequence 55316, A
5	39	62.9	591	12	US-10-282-122A-60569	Sequence 60569, A
6	39	62.9	3427	16	US-10-408-765A-1624	Sequence 1624, Ap
7	38	61.3	248	12	US-10-424-599-250800	Sequence 250800,
8	37.5	60.5	85	12	US-10-424-599-150738	Sequence 150738,
9	37	59.7	200	11	US-09-801-944B-143	Sequence 143, App
10	36	58.1	55	12	US-10-424-599-188125	Sequence 188125,
11	36	58.1	58	14	US-10-082-830-236	Sequence 236, App
12	36	58.1	73	12	US-10-424-599-247522	Sequence 247522,
13	36	58.1	94	12	US-10-424-599-199418	Sequence 199418,
14	36	58.1	322	12	US-10-282-122A-66811	Sequence 66811, A
15	36	58.1	382	9	US-09-547-267-7	Sequence 7, Appli
16	36	58.1	382	10	US-09-920-923-5	Sequence 5, Appli
17	36	58.1	382	12	US-10-695-980-5	Sequence 5, Appli
18	36	58.1	428	9	US-09-797-207-14	Sequence 14, Appl
19	36	58.1	428	12	US-10-445-318-14	Sequence 14, Appl
20	36	58.1	428	14	US-10-084-406-13	Sequence 13, Appl
21	36	58.1	428	14	US-10-388-307-13	Sequence 13, Appl
22	36	58.1	715	12	US-10-087-192-1998	Sequence 1998, Ap
23	36	58.1	748	12	US-10-087-192-1995	Sequence 1995, Ap
24	36	58.1	920	12	US-10-087-192-1518	Sequence 1518, Ap
25	36	58.1	954	16	US-10-437-963-166760	Sequence 166760,
26	36	58.1	1038	9	US-09-752-639-151	Sequence 151, App
27	36	58.1	1038	9	US-09-984-198-151	Sequence 151, App
28	36	58.1	1349	16	US-10-437-963-144453	Sequence 144453,
29	36	58.1	1953	15	US-10-369-493-1945	Sequence 1945, Ap
30	35	56.5	45	9	US-09-729-835-64	Sequence 64, Appl
31	35	56.5	45	16	US-10-373-809-64	Sequence 64, Appl
32	35	56.5	50	15	US-10-369-324-100	Sequence 100, App
33	35	56.5	50	16	US-10-607-538-125	Sequence 125, App
34	35	56.5	146	16	US-10-437-963-125125	Sequence 125125,
35	35	56.5	237	12	US-10-424-599-191896	Sequence 191896,
36	35	56.5	332	9	US-09-738-626-5674	Sequence 5674, Ap
37	35	56.5	429	16	US-10-437-963-190212	Sequence 190212,
38	35	56.5	542	12	US-10-636-396-2	Sequence 2, Appli
39	35	56.5	542	12	US-10-636-396-20	Sequence 20, Appl
40	35	56.5	542	16	US-10-636-026-2	Sequence 2, Appli
41	35	56.5	542	16	US-10-636-026-20	Sequence 20, Appl
42	35	56.5	799	10	US-09-950-041-8	Sequence 8, Appli
43	35	56.5	799	14	US-10-128-166-7	Sequence 7, Appli
44	35	56.5	799	16	US-10-732-563-8	Sequence 8, Appli
45	35	56.5	837	10	US-09-950-041-26	Sequence 26, Appl

#### ALIGNMENTS

RESULT 1

US-10-281-652-6

; Sequence 6, Application US/10281652

; Publication No. US20030091606A1

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

```

; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281,652
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-6

```

```

Query Match          100.0%; Score 62; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MPQNFYKLPQM 11
        |||||
Db      1 MPQNFYKLPQM 11

```

## RESULT 2

US-10-029-386-32258

```

; Sequence 32258, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32258
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132795.12
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
; OTHER INFORMATION: SWISSPROT HIT: Q01656, EVALUE 5.00e-21  
US-10-029-386-32258

Query Match 64.5%; Score 40; DB 14; Length 164;  
Best Local Similarity 60.0%; Pred. No. 31;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
:|::||| ||  
Db 37 LPKSFYKAPQ 46

RESULT 3

US-10-108-260A-3071

; Sequence 3071, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3071  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3071

Query Match 64.5%; Score 40; DB 15; Length 511;  
Best Local Similarity 60.0%; Pred. No. 1e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
:|::||| ||  
Db 201 LPKSFYKAPQ 210

RESULT 4

US-10-282-122A-55316

; Sequence 55316, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant

```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55316
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-282-122A-55316

```

```

Query Match          62.9%;  Score 39;  DB 12;  Length 582;
Best Local Similarity 77.8%;  Pred. No. 1.7e+02;
Matches      7;  Conservative    0;  Mismatches    2;  Indels      0;  Gaps      0;

```

```

Qy          2 PQNFYKLPQ 10
              | ||| |||
Db          190 PGNFYALPQ 198

```

```

RESULT 5
US-10-282-122A-60569
; Sequence 60569, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

```

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60569
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Listeria monocytogenes
US-10-282-122A-60569

```

```

Query Match          62.9%;   Score 39;   DB 12;   Length 591;
Best Local Similarity 77.8%;   Pred. No. 1.8e+02;
Matches      7;   Conservative    0;   Mismatches    2;   Indels      0;   Gaps      0;

```

```

Qy      2 PQNFYKLPQ 10
        | ||| |||
Db      189 PGNFYALPQ 197

```

```

RESULT 6
US-10-408-765A-1624
; Sequence 1624, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.

```

```

; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1624
; LENGTH: 3427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124,
; LOCATION: 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
; LOCATION: 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148,
; LOCATION: 149, 150, 151, 152, 153, 154, 155, 156, 157, 158
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-1624

```

```

Query Match          62.9%; Score 39; DB 16; Length 3427;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches      6; Conservative      2; Mismatches      1; Indels      0; Gaps      0;

```

```

Qy      1 MPQNFYKLP 9
        :|||||: |
Db      1235 LPQNFYEKP 1243

```

#### RESULT 7

```

US-10-424-599-250800
; Sequence 250800, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250800
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(248)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68501C.1.pep

```

US-10-424-599-250800

Query Match 61.3%; Score 38; DB 12; Length 248;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
| | : | | | |  
Db 172 PHNYYHLPQ 180

RESULT 8

US-10-424-599-150738

; Sequence 150738, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 150738

; LENGTH: 85

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(85)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_107140C.1.pep

US-10-424-599-150738

Query Match 60.5%; Score 37.5; DB 12; Length 85;  
Best Local Similarity 80.0%; Pred. No. 43;  
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 2 PQNFYK-LPQ 10  
| | | | : | | | |  
Db 27 PQNFFKELPQ 36

RESULT 9

US-09-801-944B-143

; Sequence 143, Application US/09801944B

; Publication No. US20040014169A1

; GENERAL INFORMATION:

; APPLICANT: Vogeli, Gabriel

; APPLICANT: Wood, Linda S.

; TITLE OF INVENTION: No. US20040014169A1e1 G Protein-Coupled Receptors

; FILE REFERENCE: 00100US1

```

; CURRENT APPLICATION NUMBER: US/09/801,944B
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: 60/187,828
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,715
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,929
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,930
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,825
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,833
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,830
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,829
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,582
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/187,581
; PRIOR FILING DATE: 2000-03-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 143
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-944B-143

```

```

Query Match          59.7%; Score 37; DB 11; Length 200;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      3 QNFYKLPQ 10
        |||| :||
Db      177 QNFYSVPQ 184

```

# RESULT 10

```

US-10-424-599-188125
; Sequence 188125, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 188125

```



; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140891C.1.pep  
US-10-424-599-188125

Query Match 58.1%; Score 36; DB 12; Length 55;  
Best Local Similarity 62.5%; Pred. No. 50;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQNFYKLP 9  
|:|:|:|  
Db 38 PKNFWKIP 45

RESULT 11

US-10-082-830-236  
; Sequence 236, Application US/10082830  
; Publication No. US20030077604A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Salceda, Susana  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Turner, Leah  
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific  
; TITLE OF INVENTION: Genes and Proteins  
; FILE REFERENCE: DEX-0249  
; CURRENT APPLICATION NUMBER: US/10/082,830  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,802  
; PRIOR FILING DATE: 2000-10-27  
; NUMBER OF SEQ ID NOS: 282  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 236  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-082-830-236

Query Match 58.1%; Score 36; DB 14; Length 58;  
Best Local Similarity 54.5%; Pred. No. 53;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
:|:|:|: || |  
Db 28 IPQSFFLLPAM 38

RESULT 12

US-10-424-599-247522  
; Sequence 247522, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

```
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 247522
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_65541C.1.pep
US-10-424-599-247522
```

```
Query Match          58.1%; Score 36; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 67;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 MPQNFY 6
        |||||
Db      42 MPQNFY 47
```

# RESULT 13

```
US-10-424-599-199418
; Sequence 199418, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199418
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22099C.1.pep
US-10-424-599-199418
```

```
Query Match          58.1%; Score 36; DB 12; Length 94;
Best Local Similarity 55.6%; Pred. No. 87;
Matches      5; Conservative      3; Mismatches      1; Indels      0; Gaps      0;
```

```
Qy      2 PQNFYKLPQ 10
        |: ||: ||:
Db      37 PEEFYELPE 45
```

RESULT 14  
 US-10-282-122A-66811  
 ; Sequence 66811, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 66811  
 ; LENGTH: 322  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-10-282-122A-66811

Query Match 58.1%; Score 36; DB 12; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PQNFYKLP 9

Db                   |::||:|  
232 PEDFYRLP 239

RESULT 15

US-09-547-267-7

; Sequence 7, Application US/09547267  
; Patent No. US20020147371A1  
; GENERAL INFORMATION:  
; APPLICANT: Hohmann, Hans-Peter  
; APPLICANT: Pasamontes, Luis  
; APPLICANT: Tessier, Michel  
; APPLICANT: van Loon, Adolphus  
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/547,267  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/660,645  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pokras, Bruce A.  
; REGISTRATION NUMBER: 32,748  
; REFERENCE/DOCKET NUMBER: RAN 6002/170  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201) 235-5801  
; TELEFAX: (201) 235-2363  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 382 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-547-267-7

Query Match                   58.1%;   Score 36;   DB 9;   Length 382;  
Best Local Similarity       75.0%;   Pred. No. 3.8e+02;  
Matches       6;   Conservative   1;   Mismatches   1;   Indels       0;   Gaps       0;

Qy                   3 QNFYKLPQ 10  
                     | ||:|  
Db                   330 QRFYRLPQ 337

Search completed: August 24, 2004, 16:41:22  
Job time : 41.8134 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:23:00 ; Search time 33.9851 Seconds  
(without alignments)  
102.124 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							

---

1	42	67.7	465	5	Q8I4Z3	Q8i4z3 plasmodium
2	42	67.7	474	16	Q8R8C0	Q8r8c0 thermoanaer
3	41	66.1	113	17	Q9YBW0	Q9ybw0 aeropyrum p
4	41	66.1	592	16	Q7UFY6	Q7ufy6 rhodopirell
5	40	64.5	141	6	Q27939	Q27939 alces alces
6	40	64.5	487	4	Q9NQ10	Q9nq10 homo sapien
7	40	64.5	689	5	Q9Y1I4	Q9yli4 haemonchus
8	40	64.5	837	5	O76751	O76751 haemonchus
9	39	62.9	247	16	O67490	O67490 aquifex aeo
10	39	62.9	1131	9	Q853A4	Q853a4 mycobacteri
11	39	62.9	1804	5	Q8WPD7	Q8wpd7 ciona intes
12	38	61.3	104	10	Q84TF7	Q84tf7 arabidopsis
13	38	61.3	289	10	Q9CAI8	Q9cai8 arabidopsis
14	38	61.3	445	16	Q92G77	Q92g77 rickettsia
15	38	61.3	504	5	Q86AH3	Q86ah3 dictyosteli
16	37	59.7	202	16	Q89QF7	Q89qf7 bradyrhizob
17	37	59.7	202	16	Q7WGM0	Q7wgm0 bordetella
18	37	59.7	202	16	Q7W5W8	Q7w5w8 bordetella
19	37	59.7	221	5	Q9VH11	Q9vh11 drosophila
20	37	59.7	266	16	Q55490	Q55490 synechocyst
21	37	59.7	290	5	Q23769	Q23769 chironomus
22	37	59.7	440	11	Q7TQ25	Q7tq25 rattus norv
23	37	59.7	440	11	Q7TNM3	Q7tnm3 rattus norv
24	37	59.7	491	4	Q96FT4	Q96ft4 homo sapien
25	37	59.7	656	9	Q7Y4Y1	Q7y4y1 bacteriopha
26	37	59.7	689	5	Q19491	Q19491 caenorhabdi
27	37	59.7	733	5	Q8IL63	Q8il63 plasmodium
28	37	59.7	972	5	Q26614	Q26614 strongyloce
29	37	59.7	1038	12	Q918V1	Q918v1 bovine parv
30	37	59.7	3443	11	Q8JZM8	Q8jzm8 mus musculu
31	36	58.1	209	16	Q834Z7	Q834z7 enterococcu
32	36	58.1	254	11	Q8CF26	Q8cf26 mus musculu
33	36	58.1	322	16	Q9HTV4	Q9htv4 pseudomonas
34	36	58.1	323	16	Q81RY5	Q81ry5 bacillus an
35	36	58.1	382	2	P94791	P94791 flavobacter
36	36	58.1	428	4	Q86T81	Q86t81 homo sapien
37	36	58.1	428	11	Q64165	Q64165 rattus norv
38	36	58.1	428	11	Q8BRB2	Q8brb2 mus musculu
39	36	58.1	486	5	Q19863	Q19863 caenorhabdi
40	36	58.1	564	10	Q9LNH9	Q9lnh9 arabidopsis
41	36	58.1	606	16	Q9AAN8	Q9aan8 caulobacter
42	36	58.1	638	10	Q9SRV8	Q9srv8 arabidopsis
43	36	58.1	715	4	Q8N590	Q8n590 homo sapien
44	36	58.1	715	4	Q9UNA4	Q9una4 homo sapien
45	36	58.1	715	4	Q9NYH6	Q9nyh6 homo sapien

# ALIGNMENTS

## RESULT 1

Q8I4Z3

ID Q8I4Z3 PRELIMINARY; PRT; 465 AA.  
AC Q8I4Z3;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.  
 GN PFL2100W.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3D7;  
 RX MEDLINE=22255705; PubMed=12368864;  
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
 RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,  
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
 RA Fraser C.M., Barrell B.;  
 RT "Genome sequence of the human malaria parasite Plasmodium  
 RT falciparum.";  
 RL Nature 419:498-511(2002).  
 DR EMBL; AE014851; AAN36504.1; -.  
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 465 AA; 53968 MW; 60C4D3C0DFD6E260 CRC64;

Query Match 67.7%; Score 42; DB 5; Length 465;  
 Best Local Similarity 63.6%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11  
 :|:| | | | | :  
 Db 114 IPRNFYCLPQI 124

## RESULT 2

Q8R8C0  
 ID Q8R8C0 PRELIMINARY; PRT; 474 AA.  
 AC Q8R8C0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE S-layer homology domain.  
 GN TTE2087.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;



RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of *T. tengcongensis* genome.";  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL; AE013156; AAM25260.1; -.  
 DR InterPro; IPR001119; SLH.  
 DR Pfam; PF00395; SLH; 2.  
 KW Complete proteome.  
 SQ SEQUENCE 474 AA; 53073 MW; B373A46E9E58B91E CRC64;

Query Match 67.7%; Score 42; DB 16; Length 474;  
 Best Local Similarity 77.8%; Pred. No. 14;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
 :| |||||  
 Db 260 IPTNFYKLP 268

# RESULT 3

Q9YBW0

ID Q9YBW0 PRELIMINARY; PRT; 113 AA.  
 AC Q9YBW0;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein APE1490.  
 GN APE1490.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 OC Desulfurococcaceae; Aeropyrum.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 crenarchaeon, *Aeropyrum pernix* K1.";  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000061; BAA80488.1; -.  
 DR PIR; B72629; B72629.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 113 AA; 12228 MW; 1945C293782213BE CRC64;

Query Match 66.1%; Score 41; DB 17; Length 113;  
 Best Local Similarity 54.5%; Pred. No. 5.3;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQM 11

Db                   :| ||: ||:|  
15 LPSNFFSLPRM 25

RESULT 4

Q7UFY6

ID   Q7UFY6           PRELIMINARY;       PRT;    592 AA.  
AC   Q7UFY6;  
DT   01-OCT-2003 (TrEMBLrel. 25, Created)  
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE   Aspartyl-tRNA synthetase (EC 6.1.1.12).  
GN   ASPS OR RB8253.  
OS   Rhodopirellula baltica.  
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC   Planctomycetaceae; Pirellula.  
OX   NCBI\_TaxID=117;  
RN   [1]  
RP   SEQUENCE FROM N.A.  
RC   STRAIN=1;  
RX   MEDLINE=22735913; PubMed=12835416;  
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
RA   Schlesner H., Amann R., Reinhardt R.;  
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.  
RT   strain 1.";  
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
DR   EMBL; BX294147; CAD78543.1; -.  
KW   Aminoacyl-tRNA synthetase; Ligase; Complete proteome.  
SQ   SEQUENCE   592 AA;   66228 MW;   4F1334684E8EC3AB CRC64;

Query Match                   66.1%;   Score 41;   DB 16;   Length 592;  
Best Local Similarity       77.8%;   Pred. No. 28;  
Matches       7;   Conservative       0;   Mismatches       2;   Indels       0;   Gaps       0;

Qy                   2 PQNFYKLPQ 10  
                     | ||| |||  
Db                   189 PSNFYALPQ 197

RESULT 5

Q27939

ID   Q27939           PRELIMINARY;       PRT;    141 AA.  
AC   Q27939;  
DT   01-NOV-1996 (TrEMBLrel. 01, Created)  
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE   B-casein (Fragment).  
OS   Alces alces (moose).  
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
OC   Cervidae; Odocoileinae; Alces.  
OX   NCBI\_TaxID=9852;  
RN   [1]  
RP   SEQUENCE FROM N.A.  
RX   MEDLINE=96364219; PubMed=8752004;  
RA   Gatesy J., Hayashi C., Cronin M.A., Arctander P.;

RT "Evidence from milk casein genes that cetaceans are close relatives of  
 RT hippopotamid artiodactyls.";  
 RL Mol. Biol. Evol. 13:954-963(1996).  
 DR EMBL; U53896; AAB08403.1; -.  
 DR InterPro; IPR001588; Casein.  
 DR Pfam; PF00363; caseins; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 141 141  
 SQ SEQUENCE 141 AA; 15763 MW; DC39F68595C13C72 CRC64;

Query Match 64.5%; Score 40; DB 6; Length 141;  
 Best Local Similarity 70.0%; Pred. No. 10;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 :|||| |||  
 Db 22 LPQNFLPLPQ 31

#### RESULT 6

Q9NQ10

ID Q9NQ10 PRELIMINARY; PRT; 487 AA.  
 AC Q9NQ10;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE DJ412I7.1 (Similar to radial spokehead protein) (Fragment).  
 GN DJ412I7.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews L.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL132795; CAC00771.1; -.  
 DR InterPro; IPR006802; Radial\_spoke.  
 DR Pfam; PF04712; Radial\_spoke; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 487 AA; 55796 MW; 1ADB7D60F7B5D098 CRC64;

Query Match 64.5%; Score 40; DB 4; Length 487;  
 Best Local Similarity 60.0%; Pred. No. 35;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 :|::||| ||  
 Db 177 LPKSFYKAPQ 186

#### RESULT 7

Q9Y1I4

ID Q9Y1I4 PRELIMINARY; PRT; 689 AA.  
 AC Q9Y1I4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative zinc metallopeptidase (Fragment).  
 GN MEP4.  
 OS Haemonchus contortus (Barber pole worm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.  
 OX NCBI\_TaxID=6289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Moredun;  
 RA Skuce P.J., Newlands G.F.J.;  
 RT "Molecular cloning of a putative zinc metallopeptidase (MEP4) from the  
 RT parasitic nematode Haemonchus contortus.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF132519; AAD41474.1; -.  
 DR GO; GO:0008237; F:metallopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR008753; Peptidase\_M13\_N.  
 DR Pfam; PF05649; Peptidase\_M13\_N; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 689 689  
 SQ SEQUENCE 689 AA; 78353 MW; 1431BEC11CB74E50 CRC64;

Query Match 64.5%; Score 40; DB 5; Length 689;  
 Best Local Similarity 70.0%; Pred. No. 49;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 ||: || |||  
 Db 373 MPREFYVLPQ 382

# RESULT 8

O76751  
 ID O76751 PRELIMINARY; PRT; 837 AA.  
 AC O76751;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative zinc metallopeptidase (Fragment).  
 GN MEP3.  
 OS Haemonchus contortus (Barber pole worm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.  
 OX NCBI\_TaxID=6289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Moredun;  
 RX MEDLINE=20118043; PubMed=10652118;  
 RA Smith W.D., Smith S.K., Pettit D., Newlands G.F.J., Skuce P.J.;  
 RT "Relative protective properties of three membrane glycoprotein  
 RT fractions from Haemonchus contortus.";  
 RL Parasite Immunol. 22:63-71(2000).  
 DR EMBL; AF080172; AAC31568.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008237; F:metallopeptidase activity; IEA.

DR GO; GO:0004245; F:neprilysin activity; IEA.  
 DR GO; GO:0004672; F:protein kinase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000718; Peptidase\_M13.  
 DR InterPro; IPR008753; Peptidase\_M13\_N.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF01431; Peptidase\_M13; 1.  
 DR Pfam; PF05649; Peptidase\_M13\_N; 1.  
 DR PRINTS; PR00786; NEPRILYSIN.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 837 AA; 95506 MW; 3681FA41EA66FB3E CRC64;

Query Match 64.5%; Score 40; DB 5; Length 837;  
 Best Local Similarity 70.0%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 ||: || |||  
 Db 255 MPREFYVLPQ 264

RESULT 9  
 O67490

ID O67490 PRELIMINARY; PRT; 247 AA.  
 AC O67490;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ABC transporter.  
 GN ABCT10 OR AQ\_1531.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus."  
 RL Nature 392:353-358(1998).  
 DR EMBL; AE000744; AAC07448.1; -.  
 DR PIR; H70432; H70432.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003439; ABC\_transporter.

DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KW Complete proteome.  
SQ SEQUENCE 247 AA; 27535 MW; 8C843C4C2F2AFD5F CRC64;

Query Match 62.9%; Score 39; DB 16; Length 247;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
|::||| ||  
Db 130 PEDFYKYPQ 138

RESULT 10

Q853A4

ID Q853A4 PRELIMINARY; PRT; 1131 AA.  
AC Q853A4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Gp199.  
GN 199.  
OS Mycobacteriophage Bxz1.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.  
OX NCBI\_TaxID=205877;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22592660; PubMed=12705866;  
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
RA Hatfull G.F.;  
RT "Origins of highly mosaic mycobacteriophage genomes.";  
RL Cell 113:171-182(2003).  
DR EMBL; AY129337; AAN16833.1; -.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.  
DR GO; GO:0003889; F:alpha DNA polymerase activity; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0006260; P:DNA replication; IEA.  
DR InterPro; IPR004013; PHP\_C.  
DR InterPro; IPR003141; PHP\_N.  
DR InterPro; IPR004805; PolC\_alpha.  
DR Pfam; PF02811; PHP\_C; 1.  
DR Pfam; PF02231; PHP\_N; 1.  
DR SMART; SM00481; POLIIIAC; 1.  
DR TIGRFAMs; TIGR00594; polc; 1.  
SQ SEQUENCE 1131 AA; 127013 MW; 56FA190C226F575E CRC64;

Query Match 62.9%; Score 39; DB 9; Length 1131;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
|: ||| ||:  
Db 114 PEQFYKKPQL 123

RESULT 11

Q8WPD7

ID Q8WPD7 PRELIMINARY; PRT; 1804 AA.  
AC Q8WPD7;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Complement component C3 (Fragment).  
GN C3-2.  
OS Ciona intestinalis.  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
OC Phlebobranchia; Cionidae; Ciona.  
OX NCBI\_TaxID=7719;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RA Pinto R., Marino R., Kimura Y., De Santis R., Lambris J.D.;  
RT "Isolation and characterization of two C3-like genes in Ciona  
RT intestinalis.";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ320543; CAC85958.1; -.  
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.  
DR InterPro; IPR002890; A2M\_N.  
DR InterPro; IPR009048; AM\_receptor\_bind.  
DR InterPro; IPR001599; MacrogloblnA2.  
DR InterPro; IPR001134; Netrin\_C.  
DR InterPro; IPR008930; Terp\_cyc\_toroid.  
DR Pfam; PF00207; A2M; 1.  
DR Pfam; PF01835; A2M\_N; 1.  
DR Pfam; PF01759; NTR; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 1804 AA; 201902 MW; 1BE253C037E0C5FD CRC64;

Query Match 62.9%; Score 39; DB 5; Length 1804;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLP 9  
|| |||:|  
Db 203 PQEFYKMP 210

RESULT 12

Q84TF7

ID Q84TF7 PRELIMINARY; PRT; 104 AA.  
AC Q84TF7;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE At2g42395.  
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,  
 RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis ORF clones."  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BT005839; AA064774.1; -.  
 SQ SEQUENCE 104 AA; 11914 MW; 55C7A6E407819F65 CRC64;

Query Match 61.3%; Score 38; DB 10; Length 104;  
 Best Local Similarity 66.7%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QNFYKLPQM 11  
 :||| ||:|  
 Db 49 ENFYTLPEM 57

#### RESULT 13

##### Q9CAI8

ID Q9CAI8 PRELIMINARY; PRT; 289 AA.  
 AC Q9CAI8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative DNA-binding protein.  
 GN F28P22.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,



RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*  
 RT *thaliana*.";  
 RL Nature 408:816-820(2000).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.  
 DR EMBL; AC010926; AAG51858.1; -.  
 DR PIR; C96752; C96752.  
 DR HSSP; P54274; 1BA5.  
 DR GO; GO:0000786; C:nucleosome; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006334; P:nucleosome assembly; IEA.  
 DR InterPro; IPR005818; Histone\_H1/H5.  
 DR InterPro; IPR001005; Myb\_DNA\_binding.  
 DR Pfam; PF00538; linker\_histone; 1.  
 DR Pfam; PF00249; myb\_DNA-binding; 1.  
 DR SMART; SM00526; H15; 1.  
 DR SMART; SM00717; SANT; 1.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS50090; MYB\_3; 1.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 289 AA; 32203 MW; 771BB937406EBF3C CRC64;

Query Match 61.3%; Score 38; DB 10; Length 289;  
 Best Local Similarity 85.7%; Pred. No. 49;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLP 9  
 |||||:  
 Db 186 QNFYKIP 192

#### RESULT 14

Q92G77

ID Q92G77 PRELIMINARY; PRT; 445 AA.  
 AC Q92G77;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein RC1246.  
 GN RC1246.  
 OS Rickettsia conorii.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 OC Rickettsiaceae; Rickettsieae; Rickettsia.  
 OX NCBI\_TaxID=781;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Malish 7;  
 RX MEDLINE=21442074; PubMed=11557893;  
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,  
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,  
 RA Raoult D.;  
 RT "Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*.";

RL Science 293:2093-2098(2001).  
 DR EMBL; AE008671; AAL03784.1; -.  
 DR PIR; F97855; F97855.  
 DR InterPro; IPR006638; Elp3.  
 DR InterPro; IPR006463; MiaB\_methiolase.  
 DR InterPro; IPR007197; Radical\_SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical\_SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR SMART; SM00729; Elp3; 1.  
 DR TIGRFAMs; TIGR01574; miaB-methiolase; 1.  
 DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 445 AA; 50532 MW; 0577D07F70192927 CRC64;

Query Match 61.3%; Score 38; DB 16; Length 445;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
 ||::| ||::  
 Db 109 PQSYYNLPEL 118

# RESULT 15

## Q86AH3

ID Q86AH3 PRELIMINARY; PRT; 504 AA.  
 AC Q86AH3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to Dictyostelium discoideum (Slime mold). histidine kinase  
 DE DhkK.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RX MEDLINE=22092622; PubMed=12097910;  
 RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 RA Tunggal B., Cox E., Quail M.A., Platzner M., Rosenthal A., Noegel A.A.;  
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."  
 RL Nature 418:79-85(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RA Baumgart C.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC115599; AAO51503.1; -.  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 KW Kinase.  
 SQ SEQUENCE 504 AA; 59843 MW; EAAC6E6B3C8016E7 CRC64;

Query Match 61.3%; Score 38; DB 5; Length 504;  
Best Local Similarity 60.0%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQM 11  
|| ||| |::  
Db 260 PQRFYKKPKL 269

Search completed: August 24, 2004, 15:50:48  
Job time : 37.9851 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:57:04 ; Search time 5.91045 Seconds  
(without alignments)  
96.908 Million cell updates/sec

Title: US-09-641-801-6  
Perfect score: 62  
Sequence: 1 MPQNFYKLPQM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	39	62.9	582	1	SYD_CHLMU	Q9pjk0	chlamydia m
2	39	62.9	582	1	SYD_CHLTR	O84546	chlamydia t
3	39	62.9	584	1	SYD_BUCAP	P81432	buchnera ap
4	39	62.9	591	1	SYD_LISIN	Q92bj4	listeria in
5	39	62.9	591	1	SYD_LISMO	Q8y709	listeria mo
6	39	62.9	595	1	SYD_RHIME	Q92qv4	rhizobium m
7	39	62.9	3209	1	RELN_CHICK	O93574	gallus gall
8	39	62.9	3460	1	RELN_HUMAN	P78509	homo sapien
9	39	62.9	3461	1	RELN_MOUSE	Q60841	mus musculu
10	39	62.9	3462	1	RELN_RAT	P58751	rattus norv
11	38	61.3	445	1	Y808_RICPR	Q9zce8	rickettsia
12	38	61.3	1374	1	YQ3D_SCHPO	Q09884	schizosacch
13	37	59.7	440	1	MURD_BUCAI	P57313	buchnera ap
14	37	59.7	440	1	OMGP_HUMAN	P23515	homo sapien
15	37	59.7	453	1	BRA6_CHITH	P24243	chironomus
16	37	59.7	491	1	PRCC_HUMAN	Q92733	homo sapien
17	37	59.7	609	1	SYD_CAUCR	Q9a734	caulobacter

18	37	59.7	660	1	VG06_BPT4	P19060	bacterioph
19	36	58.1	184	1	TRSF_DROSI	Q24669	drosophila
20	36	58.1	427	1	G6NT_BOVIN	Q92180	bos taurus
21	36	58.1	428	1	G6NT_HUMAN	Q02742	homo sapien
22	36	58.1	428	1	G6NT_MOUSE	Q09324	mus musculu
23	36	58.1	1953	1	BNI1_YEAST	P41832	saccharomyc
24	35	56.5	182	1	VNS1_MOUSE	Q62471	mus musculu
25	35	56.5	195	1	NUC_BORBU	O51372	borrelia bu
26	35	56.5	205	1	PTH_SYNEL	Q8dj45	synechococc
27	35	56.5	268	1	DUT_PRVKA	Q90030	pseudorabie
28	35	56.5	382	1	B4G6_MOUSE	Q9wvk5	m beta-1,4-
29	35	56.5	826	1	TLR4_PAPAN	Q9tsp2	papio anubi
30	35	56.5	839	1	TLR4_HUMAN	O00206	homo sapien
31	35	56.5	839	1	TLR4_PANPA	Q9ttn0	pan paniscu
32	34.5	55.6	527	1	YB64_YEAST	P38314	saccharomyc
33	34.5	55.6	527	1	YGF6_YEAST	P53172	saccharomyc
34	34.5	55.6	596	1	PPOB_LYCES	Q08304	lycopersico
35	34	54.8	174	1	HSCB_BUCAI	P57659	buchnera ap
36	34	54.8	202	1	SODF_METTM	Q60036	methanobact
37	34	54.8	235	1	SCE1_YEAST	P03882	saccharomyc
38	34	54.8	244	1	YWY2_CAEEL	Q11089	caenorhabdi
39	34	54.8	248	1	Y802_MYCPE	Q8euw6	mycoplasma
40	34	54.8	333	1	YC48_CYAPA	P48325	cyanophora
41	34	54.8	353	1	DDL_PROMM	Q7v819	prochloroco
42	34	54.8	446	1	Y3I0_PSEAE	Q51470	pseudomonas
43	34	54.8	474	1	YA01_PASMU	Q91699	pasteurella
44	34	54.8	474	1	YLEA_ECOLI	P77645	escherichia
45	34	54.8	474	1	YLEA_HAEIN	Q57163	haemophilus

# ALIGNMENTS

## RESULT 1

### SYD\_CHLMU

ID SYD\_CHLMU STANDARD; PRT; 582 AA.  
AC Q9PJK0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)  
DE (AspRS).  
GN ASPS OR TC0829.  
OS Chlamydia muridarum.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MoPn / Nigg;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
RA Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
RT pneumoniae AR39."

RL Nucleic Acids Res. 28:1397-1406(2000).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE002349; AAF39629.1; -.  
 DR PIR; E81660; E81660.  
 DR HSSP; P36419; 1EFW.  
 DR TIGR; TC0829; -.  
 DR HAMAP; MF\_00044; -; 1.  
 DR InterPro; IPR004524; AspS\_bact.  
 DR InterPro; IPR004115; GAD\_dom.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR004365; tRNA\_anti.  
 DR InterPro; IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF02938; GAD; 1.  
 DR Pfam; PF00152; tRNA-synt\_2; 2.  
 DR Pfam; PF01336; tRNA\_anti; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR TIGRFAMs; TIGR00459; aspS\_bact; 1.  
 DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 582 AA; 66386 MW; 492FC82B8710D420 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 582;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 190 PGNFYALPQ 198

RESULT 2  
 SYD\_CHLTR  
 ID SYD\_CHLTR STANDARD; PRT; 582 AA.  
 AC 084546;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)  
 DE (AspRS).  
 GN ASPS OR CT542.  
 OS Chlamydia trachomatis.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE001325; AAC68144.1; -.  
 DR PIR; G71500; G71500.  
 DR HSSP; P36419; 1EFW.  
 DR HAMAP; MF\_00044; -, 1.  
 DR InterPro; IPR004524; AspS\_bact.  
 DR InterPro; IPR004115; GAD\_dom.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR004365; tRNA\_anti.  
 DR InterPro; IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF02938; GAD; 1.  
 DR Pfam; PF00152; tRNA-synt\_2; 2.  
 DR Pfam; PF01336; tRNA\_anti; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR TIGRFAMs; TIGR00459; aspS\_bact; 1.  
 DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 582 AA; 66227 MW; E8D2EBB2785EFABA CRC64;

Query Match 62.9%; Score 39; DB 1; Length 582;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 190 PGNFYALPQ 198

RESULT 3

SYD\_BUCAP  
 ID SYD\_BUCAP STANDARD; PRT; 584 AA.  
 AC P81432;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)  
 DE (AspRS).  
 GN ASPS OR BUSG306.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Thao M.L., Baumann P.;  
 RT "Nucleotide sequence of a DNA fragment from Buchnera aphidicola  
 RT (Aphid endosymbiont) containing the genes aspS-trxB-serS-serC-aroA-  
 RT rpsA-himD-tpiA.";  
 RL Curr. Microbiol. 35:68-69(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22084549; PubMed=12089438;  
 RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,  
 RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;  
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";  
 RL Science 296:2376-2379(2002).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L43549; AAC05432.1; -.  
 DR EMBL; AE014107; AAM67860.1; -.  
 DR HSSP; P21889; 1EQR.  
 DR HAMAP; MF\_00044; -; 1.  
 DR InterPro; IPR004524; AspS\_bact.  
 DR InterPro; IPR004115; GAD\_dom.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR004365; tRNA\_anti.  
 DR InterPro; IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF02938; GAD; 1.  
 DR Pfam; PF00152; tRNA-synt\_2; 2.  
 DR Pfam; PF01336; tRNA\_anti; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR TIGRFAMs; TIGR00459; aspS\_bact; 1.



DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT DOMAIN 79 84 POLY-ASN.  
 SQ SEQUENCE 584 AA; 67824 MW; F318DFA1654780E7 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 584;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 182 PGNFYALPQ 190

#### RESULT 4

##### SYD LISIN

ID SYD LISIN STANDARD; PRT; 591 AA.  
 AC Q92BJ4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)  
 DE (AspRS).  
 GN ASPS OR LIN1554.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of Listeria species.";  
 RL Science 294:849-852(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

---

DR EMBL; AL596169; CAC96785.1; -.

DR PIR; AI1626; AI1626.

DR ListiList; LIN01554; -.

DR HAMAP; MF\_00044; -; 1.

DR InterPro; IPR004524; AspS\_bact.

DR InterPro; IPR004115; GAD\_dom.

DR InterPro; IPR008994; Nucleic\_acid\_OB.

DR InterPro; IPR004364; tRNA-synt\_2.

DR InterPro; IPR002312; tRNA-synt\_asp.

DR InterPro; IPR004365; tRNA\_anti.

DR InterPro; IPR006195; tRNA\_ligase\_II.

DR Pfam; PF02938; GAD; 1.

DR Pfam; PF00152; tRNA-synt\_2; 2.

DR Pfam; PF01336; tRNA\_anti; 1.

DR PRINTS; PR01042; TRNASYNTHASP.

DR TIGRFAMS; TIGR00459; aspS\_bact; 1.

DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW Complete proteome.

SQ SEQUENCE 591 AA; 66374 MW; 81AEAC0D02147F7E CRC64;

Query Match 62.9%; Score 39; DB 1; Length 591;

Best Local Similarity 77.8%; Pred. No. 12;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PQNFYKLPQ 10

| ||| |||

Db 189 PGNFYALPQ 197

# RESULT 5

## SYD\_LISMO

ID SYD\_LISMO STANDARD; PRT; 591 AA.

AC Q8Y709;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)

DE (AspRS).

GN ASPS OR LMO1519.

OS *Listeria monocytogenes*.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; *Listeria*.

OX NCBI\_TaxID=1639;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

RX MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria* species.";  
 RL Science 294:849-852(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL591979; CAC99597.1; -.  
 DR PIR; AG1264; AG1264.  
 DR ListiList; LMO01519; -.  
 DR HAMAP; MF\_00044; -; 1.  
 DR InterPro; IPR004524; AspS\_bact.  
 DR InterPro; IPR004115; GAD\_dom.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR004365; tRNA\_anti.  
 DR InterPro; IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF02938; GAD; 1.  
 DR Pfam; PF00152; tRNA-synt\_2; 2.  
 DR Pfam; PF01336; tRNA\_anti; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR TIGRFAMS; TIGR00459; aspS\_bact; 1.  
 DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 591 AA; 66382 MW; DE143FD009106BBA CRC64;

Query Match 62.9%; Score 39; DB 1; Length 591;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||  
 Db 189 PGNFYALPQ 197

RESULT 6  
 SYD\_RHIME  
 ID SYD\_RHIME STANDARD; PRT; 595 AA.  
 AC Q92QV4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)

DE (AspRS).  
 GN ASPS OR R01194 OR SMC01756.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +  
 CC diphosphate + L-aspartyl-tRNA(Asp).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL591786; CAC45773.1; -.  
 DR HAMAP; MF\_00044; -; 1.  
 DR InterPro; IPR004524; AspS\_bact.  
 DR InterPro; IPR004115; GAD\_dom.  
 DR InterPro; IPR008994; Nucleic\_acid\_OB.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR004365; tRNA\_anti.  
 DR InterPro; IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF02938; GAD; 1.  
 DR Pfam; PF00152; tRNA-synt\_2; 2.  
 DR Pfam; PF01336; tRNA\_anti; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR TIGRFAMs; TIGR00459; aspS\_bact; 1.  
 DR PROSITE; PS50862; AA\_TRNA\_LIGASE\_II; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 595 AA; 67189 MW; 59180C1B5E637BD8 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 595;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| |||

## RESULT 7

## RELN\_CHICK

ID RELN\_CHICK STANDARD; PRT; 3209 AA.

AC 093574;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Reelin (EC 3.4.21.-) (Fragment).

GN RELN.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20320899; PubMed=10861519;

RA Bernier B., Bar I., D'Arcangelo G., Curran T., Goffinet A.M.;

RT "Reelin mRNA expression during embryonic brain development in the chick.";

RL J. Comp. Neurol. 422:448-463(2000).

CC -!- FUNCTION: Extracellular matrix serine protease that plays a role  
CC in layering of neurons in the cerebral cortex and cerebellum.  
CC Regulates microtubule function in neurons and neuronal migration.  
CC Affects migration of sympathetic preganglionic neurons in the  
CC spinal cord, where it seems to act as a barrier to neuronal  
CC migration. Enzymatic activity is important for the modulation of  
CC cell adhesion. Binding to the extracellular domains of lipoprotein  
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of  
CC Dab1 and modulation of Tau phosphorylation (By similarity).

CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By  
CC similarity).

CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By  
CC similarity).

CC -!- SIMILARITY: Belongs to the reelin family.

CC -!- SIMILARITY: Contains 8 EGF-like domains.

CC -!- SIMILARITY: Contains 15 BNR repeats.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AF090441; AAC35559.1; -.

DR HSSP; P05106; 1JV2.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR002860; GH\_BNR.

DR InterPro; IPR006210; IEGF.

DR Pfam; PF02012; BNR; 15.

DR Pfam; PF00008; EGF; 3.

DR SMART; SM00181; EGF; 5.  
 DR PROSITE; PS00022; EGF\_1; 7.  
 DR PROSITE; PS01186; EGF\_2; 7.  
 DR PROSITE; PS50026; EGF\_3; 3.  
 KW Hydrolase; Serine protease; Developmental protein; Matrix protein;  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat.  
 FT NON\_TER 1 1  
 FT DOMAIN 418 449 EGF-LIKE 1.  
 FT DOMAIN 777 808 EGF-LIKE 2.  
 FT DOMAIN 1157 1190 EGF-LIKE 3.  
 FT DOMAIN 1513 1544 EGF-LIKE 4.  
 FT DOMAIN 1877 1909 EGF-LIKE 5.  
 FT DOMAIN 2226 2257 EGF-LIKE 6.  
 FT DOMAIN 2601 2632 EGF-LIKE 7.  
 FT DOMAIN 2976 3008 EGF-LIKE 8.  
 FT REPEAT 340 351 BNR 1.  
 FT REPEAT 546 557 BNR 2.  
 FT REPEAT 699 710 BNR 3.  
 FT REPEAT 904 915 BNR 4.  
 FT REPEAT 1070 1081 BNR 5.  
 FT REPEAT 1283 1294 BNR 6.  
 FT REPEAT 1434 1445 BNR 7.  
 FT REPEAT 1632 1643 BNR 8.  
 FT REPEAT 1791 1802 BNR 9.  
 FT REPEAT 1998 2009 BNR 10.  
 FT REPEAT 2147 2158 BNR 11.  
 FT REPEAT 2346 2357 BNR 12.  
 FT REPEAT 2526 2537 BNR 13.  
 FT REPEAT 2727 2738 BNR 14.  
 FT REPEAT 3111 3122 BNR 15.  
 FT DOMAIN 3180 3209 ARG-RICH (BASIC).  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1195 1195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1669 1669 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1893 1893 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2017 2017 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2065 2065 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2317 2317 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2710 2710 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2764 2764 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2821 2821 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2933 2933 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3160 3160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3187 3187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 3209 AA; 361288 MW; 81A7B6676BCAA3D1 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 3209;  
 Best Local Similarity 66.7%; Pred. No. 73;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
 :|||||: |

RESULT 8

RELN\_HUMAN

ID RELN\_HUMAN STANDARD; PRT; 3460 AA.  
AC P78509; Q9UDQ2;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Reelin precursor (EC 3.4.21.-).  
GN RELN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=97202106; PubMed=9049633;  
RA DeSilva U., D'Arcangelo G., Braden V.V., Chen J., Miao G.G.,  
RA Curran T., Green E.D.;  
RT "The human reelin gene: isolation, sequencing, and mapping on  
RT chromosome 7.";  
RL Genome Res. 7:157-164(1997).  
RN [2]  
RP SEQUENCE OF 194-2556 FROM N.A.  
RA Lamar B., Wamsley P., Gibson A., Maas J., Bauer C., Sapetti L.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP ALTERNATIVE SPLICING.  
RX MEDLINE=99263436; PubMed=10328932;  
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergeyck V.,  
RA Goffinet A.M.;  
RT "Evolutionarily conserved, alternative splicing of reelin during brain  
RT development.";  
RL Exp. Neurol. 156:229-238(1999).  
RN [4]  
RP DISEASE, AND TISSUE SPECIFICITY.  
RX MEDLINE=99080080; PubMed=9861036;  
RA Impagnatiello F., Guidotti A.R., Pesold C., Dwivedi Y., Caruncho H.,  
RA Pisu M.G., Uzunov D.P., Smalheiser N.R., Davis J.M., Pandey G.N.,  
RA Pappas G.D., Tueting P., Sharma R.P., Costa E.;  
RT "A decrease of reelin expression as a putative vulnerability factor in  
RT schizophrenia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:15718-15723(1998).  
RN [5]  
RP DISEASE.  
RX MEDLINE=20428190; PubMed=10973257;  
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahwan S.A., Grant P.E.,  
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.;  
RT "Autosomal recessive lissencephaly with cerebellar hypoplasia is  
RT associated with human RELN mutations.";  
RL Nat. Genet. 26:93-96(2000).  
RN [6]  
RP ERRATUM.  
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahwan S.A., Grant P.E.,  
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.;

RL Nat. Genet. 27:225-225(2001).  
 RN [7]  
 RP DISEASE.  
 RX MEDLINE=21217116; PubMed=11317216;  
 RA Persico A.M., D'Agruma L., Maiorano N., Totaro A., Militeri R.,  
 RA Bravaccio C., Wassink T.H., Schneider C., Melmed R., Trillo S.,  
 RA Montecchi F., Palermo M., Pascucci T., Puglisi-Allegra S.,  
 RA Reichelt K.-L., Conciatori M., Marino R., Quattrocchi C.C., Baldi A.,  
 RA Zelante L., Gasparini P., Keller F.;  
 RT "Reelin gene alleles and haplotypes as a factor predisposing to  
 RT autistic disorder.";  
 RL Mol. Psych. 6:150-159(2001).  
 CC -!- FUNCTION: Extracellular matrix serine protease that plays a role  
 CC in layering of neurons in the cerebral cortex and cerebellum.  
 CC Regulates microtubule function in neurons and neuronal migration.  
 CC Affects migration of sympathetic preganglionic neurons in the  
 CC spinal cord, where it seems to act as a barrier to neuronal  
 CC migration. Enzymatic activity is important for the modulation of  
 CC cell adhesion. Binding to the extracellular domains of lipoprotein  
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of  
 CC Dab1 and modulation of Tau phosphorylation (By similarity).  
 CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=P78509-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P78509-2; Sequence=VSP\_005575;  
 CC Name=3;  
 CC IsoId=P78509-3; Sequence=VSP\_005576;  
 CC -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis  
 CC by the Cajal-Retzius cells and other pioneer neurons located in  
 CC the telencephalic marginal zone and by granule cells of the  
 CC external granular layer of the cerebellum. In adult brain,  
 CC preferentially expressed in GABAergic interneurons of prefrontal  
 CC cortices, temporal cortex, hippocampus and glutamatergic granule  
 CC cells of cerebellum. Also expressed in fetal and adult liver.  
 CC -!- DEVELOPMENTAL STAGE: Expressed in fetal and postnatal brain and  
 CC liver. Expression in postnatal human brain is high in the  
 CC cerebellum.  
 CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By  
 CC similarity).  
 CC -!- DISEASE: Defects in RELN are the cause of autosomal recessive  
 CC lissencephaly with cerebellar hypoplasia [MIM:257320]; also known  
 CC as Norman-Roberts syndrome. Some patients also displayed  
 CC persistent lymphedema neonatally, and one showed accumulation of  
 CC chylous or fatty, ascites fluid.  
 CC -!- DISEASE: Defects in RELN may contribute to susceptibility to  
 CC schizophrenia. Expression of the protein is reduced to about 50%  
 CC in patients with schizophrenia.  
 CC -!- DISEASE: Defects in RELN may predispose to autistic disorder. A  
 CC polymorphic GGC triplet repeat located in the 5'UTR region of RELN  
 CC gene, which harbors in the normal population 8 to 10 repeats, is  
 CC significantly increased in autistic patients to carry 4 to 23  
 CC additional repeats.



CC -!- SIMILARITY: Belongs to the reelin family.  
CC -!- SIMILARITY: Contains 8 EGF-like domains.  
CC -!- SIMILARITY: Contains 15 BNR repeats.

---

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

---

DR EMBL; U79716; AAC51105.1; -.  
DR EMBL; AC000121; AAB46357.1; -.  
DR EMBL; AC006316; AAD29127.1; -.  
DR HSSP; P05106; 1JV2.  
DR Genew; HGNC:9957; RELN.  
DR MIM; 600514; -.  
DR MIM; 257320; -.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002860; GH\_BNR.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002861; Reeler.  
DR Pfam; PF02012; BNR; 15.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF02014; Reeler; 1.  
DR SMART; SM00181; EGF; 5.  
DR PROSITE; PS00022; EGF\_1; 7.  
DR PROSITE; PS01186; EGF\_2; 6.  
DR PROSITE; PS50026; EGF\_3; 5.  
KW Hydrolase; Serine protease; Developmental protein; Matrix protein;  
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;  
KW Alternative splicing; Lissencephaly.

FT	SIGNAL	1	25	POTENTIAL.
FT	CHAIN	26	3460	REELIN.
FT	DOMAIN	39	171	REELER.
FT	DOMAIN	670	701	EGF-LIKE 1.
FT	DOMAIN	1029	1060	EGF-LIKE 2.
FT	DOMAIN	1408	1441	EGF-LIKE 3.
FT	DOMAIN	1764	1795	EGF-LIKE 4.
FT	DOMAIN	2128	2160	EGF-LIKE 5.
FT	DOMAIN	2477	2508	EGF-LIKE 6.
FT	DOMAIN	2852	2883	EGF-LIKE 7.
FT	DOMAIN	3227	3259	EGF-LIKE 8.
FT	REPEAT	592	603	BNR 1.
FT	REPEAT	798	809	BNR 2.
FT	REPEAT	951	962	BNR 3.
FT	REPEAT	1156	1167	BNR 4.
FT	REPEAT	1322	1333	BNR 5.
FT	REPEAT	1534	1545	BNR 6.
FT	REPEAT	1685	1696	BNR 7.
FT	REPEAT	1883	1894	BNR 8.
FT	REPEAT	2042	2053	BNR 9.
FT	REPEAT	2249	2260	BNR 10.
FT	REPEAT	2398	2409	BNR 11.
FT	REPEAT	2597	2608	BNR 12.
FT	REPEAT	2777	2788	BNR 13.

FT	REPEAT	2978	2989	BNR 14.
FT	REPEAT	3362	3373	BNR 15.
FT	DOMAIN	3431	3460	ARG-RICH (BASIC).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	257	257	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	305	305	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	628	628	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1266	1266	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1599	1599	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1749	1749	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1920	1920	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2144	2144	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2268	2268	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2316	2316	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2568	2568	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2961	2961	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3015	3015	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3072	3072	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3184	3184	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3411	3411	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3438	3438	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	3428	3429	Missing (in isoform 2).
FT				/FTId=VSP_005575.
FT	VARSPLIC	3428	3460	Missing (in isoform 3).
FT				/FTId=VSP_005576.
FT	CONFLICT	752	752	E -> D (IN REF. 2).
SQ	SEQUENCE	3460 AA;	388399 MW;	BB2C0BA415AB93C1 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 3460;  
 Best Local Similarity 66.7%; Pred. No. 79;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
 :||||: |  
 Db 1235 LPQNFYEKP 1243

# RESULT 9

## RELN\_MOUSE

ID RELN\_MOUSE STANDARD; PRT; 3461 AA.  
 AC Q60841; Q9CUA6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Reelin precursor (EC 3.4.21.-) (Reeler protein).  
 GN RELN OR RL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=95231649; PubMed=7715726;  
 RA D'Arcangelo G., Miao G.G., Chen S.-C., Soares H.D., Morgan J.I.,  
 RA Curran T.;

RT "A protein related to extracellular matrix proteins deleted in the  
 RT mouse mutant reeler.";  
 RL Nature 374:719-723(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=98086481; PubMed=9417911;  
 RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demirov D.,  
 RA Goffinet A.M.;  
 RT "Genomic organization of the mouse reelin gene.";  
 RL Genomics 46:240-250(1997).  
 RN [3]  
 RP SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=95375789; PubMed=7647795;  
 RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,  
 RA Ohashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,  
 RA Nakao K., Katsuki M., Hayashizaki Y.;  
 RT "The reeler gene encodes a protein with an EGF-like motif expressed by  
 RT pioneer neurons.";  
 RL Nat. Genet. 10:77-83(1995).  
 RN [4]  
 RP SEQUENCE OF 3044-3461 FROM N.A. (ISOFORM 2).  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=97141547; PubMed=8987733;  
 RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,  
 RA Curran T.;  
 RT "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal  
 RT antibody.";  
 RL J. Neurosci. 17:23-31(1997).  
 RN [6]  
 RP CHARACTERIZATION.  
 RX MEDLINE=21634904; PubMed=11689558;  
 RA Quattrocchi C.C., Wannenes F., Persico A.M., Ciafre S.A.,  
 RA D'Arcangelo G., Farace M.G., Keller F.;

RT "Reelin is a serine protease of the extracellular matrix.";
 RL J. Biol. Chem. 277:303-309(2002).
 RN [7]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=97325946; PubMed=9182958;
 RA Schiffmann S.N., Bernier B., Goffinet A.M.;
 RT "Reelin mRNA expression during mouse brain development.";
 RL Eur. J. Neurosci. 9:1055-1071(1997).
 RN [8]
 RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX MEDLINE=99263436; PubMed=10328932;
 RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergeyck V.,
 RA Goffinet A.M.;
 RT "Evolutionarily conserved, alternative splicing of reelin during brain
 RT development.";
 RL Exp. Neurol. 156:229-238(1999).
 RN [9]
 RP BINDING TO VLDLR AND APOER2.
 RX MEDLINE=20036019; PubMed=10571241;
 RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
 RA Cooper J.A., Herz J.;
 RT "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
 RT tyrosine phosphorylation of disabled-1 and modulates tau
 RT phosphorylation.";
 RL Neuron 24:481-489(1999).
 RN [10]
 RP FUNCTION.
 RX MEDLINE=20359755; PubMed=10880573;
 RA Yip J.W., Yip Y.P.L., Nakajima K., Capriotti C.;
 RT "Reelin controls position of autonomic neurons in the spinal cord.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).
 CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
 CC in layering of neurons in the cerebral cortex and cerebellum.
 CC Regulates microtubule function in neurons and neuronal migration.
 CC Affects migration of sympathetic preganglionic neurons in the
 CC spinal cord, where it seems to act as a barrier to neuronal
 CC migration. Enzymatic activity is important for the modulation of
 CC cell adhesion. Binding to the extracellular domains of lipoprotein
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
 CC Dab1 and modulation of Tau phosphorylation.
 CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q60841-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q60841-2; Sequence=VSP\_005577;
 CC Name=3;
 CC IsoId=Q60841-3; Sequence=VSP\_005578;
 CC -!- TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is
 CC abundantly produced during brain ontogenesis by the Cajal-Retzius
 CC cells and other pioneer neurons located in the telencephalic
 CC marginal zone and by granule cells of the external granular layer
 CC of the cerebellum. Expression is located in deeper layers in the
 CC developing hippocampus and olfactory bulb, low levels of
 CC expression are also detected in the immature striatum. At early

developmental stages, expressed also in hypothalamic differentiation fields, tectum and spinal cord. A moderate to low level of expression occurs in the septal area, striatal fields, habenular nuclei, some thalamic nuclei, particularly the lateral geniculate, the retina and some nuclei of the reticular formation in the central field of the medulla. Very low levels found in liver and kidney. No expression in radial glial cells, cortical plate, Purkinje cells and inferior olivary neurons. The minor isoform 2 is only expressed in non neuronal cells. The minor isoform 3 is found in the same cells as isoform 1, but is almost undetectable in retina and brain stem.

CC    -!- DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.  
 CC       Expression increases up to birth and remains high from post-natal  
 CC       day 2 to 11 in both cerebellum and fore/midbrain. Expression  
 CC       declines thereafter and is largely brain specific in the adult.

CC    -!- DOMAIN: The basic C-terminal region is essential for secretion.

CC    -!- PTM: N-glycosylated and to a lesser extent also O-glycosylated.

CC    -!- DISEASE: Defects in reln are the cause of the autosomal recessive  
 CC       reeler (rl) phenotype which is characterized by impaired motor  
 CC       coordination, tremors and ataxia. Neurons in affected mice fail to  
 CC       reach their correct locations in the developing brain, disrupting  
 CC       the organization of the cerebellar and cerebral cortices and other  
 CC       laminated regions.

CC    -!- SIMILARITY: Belongs to the reelin family.

CC    -!- SIMILARITY: Contains 8 EGF-like domains.

CC    -!- SIMILARITY: Contains 15 BNR repeats.

CC    -----  
 CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC    -----

DR    EMBL; U24703; AAB91599.1; -.

DR    EMBL; D63520; BAA09788.1; ALT\_INIT.

DR    EMBL; AK017094; BAB30592.1; -.

DR    MGD; MGI:103022; Reln.

DR    GO; GO:0005615; C:extracellular space; IDA.

DR    GO; GO:0007420; P:brain development; IMP.

DR    GO; GO:0016477; P:cell migration; IMP.

DR    InterPro; IPR006209; EGF\_like.

DR    InterPro; IPR002860; GH\_BNR.

DR    InterPro; IPR006210; IEGF.

DR    InterPro; IPR002861; Reeler.

DR    Pfam; PF02012; BNR; 15.

DR    Pfam; PF00008; EGF; 3.

DR    Pfam; PF02014; Reeler; 1.

DR    SMART; SM00181; EGF; 5.

DR    PROSITE; PS00022; EGF\_1; 7.

DR    PROSITE; PS01186; EGF\_2; 6.

DR    PROSITE; PS50026; EGF\_3; 5.

KW    Hydrolase; Serine protease; Developmental protein; Matrix protein;  
 KW    Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;  
 KW    Alternative splicing.

FT    SIGNAL           1       26           POTENTIAL.

FT	CHAIN	27	3461	REELIN.
FT	DOMAIN	40	172	REELER.
FT	DOMAIN	671	702	EGF-LIKE 1.
FT	DOMAIN	1030	1061	EGF-LIKE 2.
FT	DOMAIN	1409	1442	EGF-LIKE 3.
FT	DOMAIN	1765	1796	EGF-LIKE 4.
FT	DOMAIN	2129	2161	EGF-LIKE 5.
FT	DOMAIN	2478	2509	EGF-LIKE 6.
FT	DOMAIN	2853	2884	EGF-LIKE 7.
FT	DOMAIN	3228	3260	EGF-LIKE 8.
FT	REPEAT	593	604	BNR 1.
FT	REPEAT	799	810	BNR 2.

Query Match 62.9%; Score 39; DB 1; Length 3461;  
 Best Local Similarity 66.7%; Pred. No. 79;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
 :||||: |  
 Db 1236 LPQNFYEKP 1244

# RESULT 10

## RELN\_RAT

ID RELN\_RAT STANDARD; PRT; 3462 AA.  
 AC P58751; Q80T65;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Reelin precursor (EC 3.4.21.-).  
 GN RELN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Cerebellum;  
 RA Kikkawa S., Terashima T.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND DISEASE.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=22557166; PubMed=12670697;  
 RA Yokoi N., Namae M., Wang H.-W., Kojima K., Fuse M., Yasuda K.,  
 RA Serikawa T., Seino S., Komeda K.;  
 RT "Rat neurological disease creeping is caused by a mutation in the  
 RT reelin gene.";  
 RL Brain Res. Mol. Brain Res. 112:1-7(2003).  
 RN [3]  
 RP ALTERNATIVE SPLICING.  
 RX MEDLINE=99263436; PubMed=10328932;  
 RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueyck V.,  
 RA Goffinet A.M.;  
 RT "Evolutionarily conserved, alternative splicing of reelin during brain  
 RT development.";  
 RL Exp. Neurol. 156:229-238(1999).

CC    -!- FUNCTION: Extracellular matrix serine protease that plays a role  
 CC           in layering of neurons in the cerebral cortex and cerebellum.  
 CC           Regulates microtubule function in neurons and neuronal migration.  
 CC           Affects migration of sympathetic preganglionic neurons in the  
 CC           spinal cord, where it seems to act as a barrier to neuronal  
 CC           migration. Enzymatic activity is important for the modulation of  
 CC           cell adhesion. Binding to the extracellular domains of lipoprotein  
 CC           receptors VLDLR and ApoER2 induces tyrosine phosphorylation of  
 CC           Dab1 and modulation of Tau phosphorylation (By similarity).  
 CC    -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By  
 CC           similarity).  
 CC    -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC    -!- ALTERNATIVE PRODUCTS:  
 CC           Event=Alternative splicing; Named isoforms=3;  
 CC           Name=1;  
 CC                IsoId=P58751-1; Sequence=Displayed;  
 CC           Name=2;  
 CC                IsoId=P58751-2; Sequence=VSP\_005579;  
 CC           Name=3;  
 CC                IsoId=P58751-3; Sequence=VSP\_005580;  
 CC    -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis  
 CC           by the Cajal-Retzius cells and other pioneer neurons located in  
 CC           the telencephalic marginal zone and by granule cells of the  
 CC           external granular layer of the cerebellum.  
 CC    -!- DOMAIN: The basic C-terminal region is essential for secretion (By  
 CC           similarity).  
 CC    -!- DISEASE: Defects in Reln are the cause of creeping, which is  
 CC           characterized by tremor, gait ataxia, cerebellar hypoplasia and  
 CC           abnormal neuronal migration (particularly in the cerebral cortex  
 CC           and hippocampus). The mutation is due to a nucleotide insertion at  
 CC           codon 1892 which results in a translational frameshift and  
 CC           truncation of the protein.  
 CC    -!- SIMILARITY: Belongs to the reelin family.  
 CC    -!- SIMILARITY: Contains 8 EGF-like domains.  
 CC    -!- SIMILARITY: Contains 15 BNR repeats.

CC    -----  
 CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC    -----

DR    EMBL; AB049473; BAB78470.1; -.  
 DR    EMBL; AB062680; BAC75467.1; -.  
 DR    InterPro; IPR006209; EGF\_like.  
 DR    InterPro; IPR002860; GH\_BNR.  
 DR    InterPro; IPR006210; IEGF.  
 DR    InterPro; IPR002861; Reeler.  
 DR    Pfam; PF02012; BNR; 15.  
 DR    Pfam; PF00008; EGF; 3.  
 DR    Pfam; PF02014; Reeler; 1.  
 DR    SMART; SM00181; EGF; 6.  
 DR    PROSITE; PS00022; EGF\_1; 7.  
 DR    PROSITE; PS01186; EGF\_2; 6.  
 DR    PROSITE; PS50026; EGF\_3; 5.

KW Hydrolase; Serine protease; Developmental protein; Matrix protein;  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;  
 KW Alternative splicing.

FT	SIGNAL	1	27	POTENTIAL.
FT	CHAIN	28	3462	REELIN.
FT	DOMAIN	41	173	REELER.
FT	DOMAIN	672	703	EGF-LIKE 1.
FT	DOMAIN	1031	1062	EGF-LIKE 2.
FT	DOMAIN	1410	1443	EGF-LIKE 3.
FT	DOMAIN	1766	1797	EGF-LIKE 4.
FT	DOMAIN	2130	2162	EGF-LIKE 5.
FT	DOMAIN	2479	2510	EGF-LIKE 6.
FT	DOMAIN	2854	2885	EGF-LIKE 7.
FT	DOMAIN	3229	3261	EGF-LIKE 8.
FT	REPEAT	594	605	BNR 1.
FT	REPEAT	800	811	BNR 2.
FT	REPEAT	953	964	BNR 3.
FT	REPEAT	1158	1169	BNR 4.
FT	REPEAT	1324	1335	BNR 5.
FT	REPEAT	1536	1547	BNR 6.
FT	REPEAT	1687	1698	BNR 7.
FT	REPEAT	1885	1896	BNR 8.
FT	REPEAT	2044	2055	BNR 9.
FT	REPEAT	2251	2262	BNR 10.
FT	REPEAT	2400	2411	BNR 11.
FT	REPEAT	2599	2610	BNR 12.
FT	REPEAT	2779	2790	BNR 13.
FT	REPEAT	2980	2991	BNR 14.
FT	REPEAT	3364	3375	BNR 15.
FT	DOMAIN	3433	3462	ARG-RICH (BASIC).
FT	CARBOHYD	142	142	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	291	291	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	307	307	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	630	630	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1268	1268	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1448	1448	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1601	1601	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1751	1751	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1922	1922	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2146	2146	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2270	2270	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2318	2318	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2570	2570	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2963	2963	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3017	3017	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3074	3074	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3186	3186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3413	3413	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3440	3440	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	3430	3431	Missing (in isoform 2).
FT				/FTId=VSP_005579.
FT	VARSPLIC	3430	3462	Missing (in isoform 3).
FT				/FTId=VSP_005580.
FT	CONFLICT	336	336	H -> R (IN REF. 2).
FT	CONFLICT	2714	2714	V -> L (IN REF. 2).
SQ	SEQUENCE	3462	AA; 387525	MW; FCCF89B090E035F6 CRC64;



Query Match 62.9%; Score 39; DB 1; Length 3462;  
Best Local Similarity 66.7%; Pred. No. 79;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPQNFYKLP 9  
:||||: |  
Db 1237 LPQNFYEKP 1245

RESULT 11

Y808\_RICPR

ID Y808\_RICPR STANDARD; PRT; 445 AA.  
AC Q9ZCE8;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein RP808.  
GN RP808.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Rickettsieae; Rickettsia.  
OX NCBI\_TaxID=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=99039499; PubMed=9823893;  
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
RT mitochondria.";  
RL Nature 396:133-140(1998).  
CC -!- SIMILARITY: Belongs to the UPF0004 family.  
CC -!- SIMILARITY: Contains 1 TRAM domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ235273; CAA15234.1; -.  
DR PIR; B71642; B71642.  
DR InterPro; IPR006638; Elp3.  
DR InterPro; IPR006463; MiaB\_methiolase.  
DR InterPro; IPR007197; Radical\_SAM.  
DR InterPro; IPR002792; TRAM.  
DR InterPro; IPR005839; UPF0004.  
DR Pfam; PF04055; Radical\_SAM; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR Pfam; PF00919; UPF0004; 1.  
DR SMART; SM00729; Elp3; 1.  
DR TIGRFAMs; TIGR01574; miaB-methiolase; 1.  
DR TIGRFAMs; TIGR00089; TIGR00089; 1.

DR PROSITE; PS50926; TRAM; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 FT DOMAIN 383 445 TRAM.  
 SQ SEQUENCE 445 AA; 50540 MW; 31246EF79EF613B2 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 445;  
 Best Local Similarity 50.0%; Pred. No. 14;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PQNFYKLPQM 11  
 ||::| ||::  
 Db 109 PQSYYNLPEL 118

RESULT 12  
 YQ3D\_SCHPO

ID YQ3D\_SCHPO STANDARD; PRT; 1374 AA.  
 AC Q09884; Q9UUN1;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative helicase Cl88.13c (EC 3.6.1.-).  
 GN SPCC188.13C OR SPCC584.10C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- SIMILARITY: Belongs to the helicase family.  
 CC -!- SIMILARITY: Contains 2 RNase III domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL032824; CAB37423.1; -.  
 DR EMBL; AL049662; CAB41233.1; -.  
 DR PIR; T39130; S62524.  
 DR GeneDB\_SPombe; SPCC188.13c; -.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR005034; DUF283.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000999; RNase\_III.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF03368; DUF283; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00636; Ribonuclease\_3; 2.  
 DR SMART; SM00487; DEXDc; 1.  
 DR SMART; SM00490; HELICc; 1.  
 DR SMART; SM00535; RIBOc; 2.  
 DR PROSITE; PS00517; RNASE\_3\_1; 1.  
 DR PROSITE; PS50142; RNASE\_3\_2; 2.  
 KW Hypothetical protein; Helicase; ATP-binding; Hydrolase; Nuclease;  
 KW Endonuclease; Repeat.  
 FT NP\_BIND 32 39 ATP (POTENTIAL).  
 FT SITE 145 148 DECH BOX.  
 FT DOMAIN 916 1038 RNASE III 1.  
 FT DOMAIN 1083 1233 RNASE III 2.  
 SQ SEQUENCE 1374 AA; 158039 MW; 89AE9EF8DE7966C6 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 1374;  
 Best Local Similarity 75.0%; Pred. No. 45;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLPQ 10  
 |||:|  
 Db 779 QNFFKLPE 786

# RESULT 13

## MURD\_BUCAI

ID MURD\_BUCAI STANDARD; PRT; 440 AA.  
 AC P57313;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9) (UDP-N-  
 DE acetylmuramoyl-L-alanyl-D-glutamate synthetase) (D-glutamic acid  
 DE adding enzyme).

GN MURD OR BU218.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tokyo 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC -!- FUNCTION: Cell wall formation. Catalyzes the addition of  
 CC glutamate to the nucleotide precursor UDP-N-acetylmuramoyl-L-  
 CC alanine (UMA) (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanine +  
 CC glutamate = ADP + phosphate + UDP-N-acetylmuramoyl-L-alanyl-D-  
 CC glutamate.  
 CC -!- PATHWAY: Peptidoglycan biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the murCDEF family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP001118; BAB12934.1; -.  
 DR HSSP; P14900; 1E0D.  
 DR HAMAP; MF\_00639; -, 1.  
 DR InterPro; IPR000713; Mur\_ligase.  
 DR InterPro; IPR004101; Mur\_ligase\_C.  
 DR InterPro; IPR005762; MurD.  
 DR Pfam; PF01225; Mur\_ligase; 1.  
 DR Pfam; PF02875; Mur\_ligase\_C; 1.  
 DR TIGRFAMs; TIGR01087; murD; 1.  
 KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;  
 KW ATP-binding; Complete proteome.  
 FT NP\_BIND 113 119 ATP (POTENTIAL).  
 SQ SEQUENCE 440 AA; 49528 MW; B343A7235FD9AE33 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 440;  
 Best Local Similarity 66.7%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PQNFYKLPQ 10  
 | ||| :||  
 Db 42 PSNFIKIPQ 50

RESULT 14  
 OMGP\_HUMAN

ID OMGP\_HUMAN STANDARD; PRT; 440 AA.  
 AC P23515;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Oligodendrocyte-myelin glycoprotein precursor.  
 GN OMG OR OMGP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91117257; PubMed=1899288;  
 RA Viskochil D., Cawthon R.M., O'Connell P., Xu G., Stevens J.,  
 RA Culver M., Carey J., White R.;  
 RT "The gene encoding the oligodendrocyte-myelin glycoprotein is  
 RT embedded within the neurofibromatosis type 1 gene.";  
 RL Mol. Cell. Biol. 11:906-912(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90130636; PubMed=1688857;  
 RA Mikol D.D., Gulcher J., Stefansson K.;  
 RT "The oligodendrocyte-myelin glycoprotein belongs to a distinct family  
 RT of proteins and contains the HNK-1 carbohydrate.";  
 RL J. Cell Biol. 110:471-479(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91115958; PubMed=2277079;  
 RA Mikol D.D., Alexakos M.J., Bayley C.A., Lemons R.S., le Beau M.M.,  
 RA Stefansson K.;  
 RT "Structure and chromosomal localization of the gene for the  
 RT oligodendrocyte-myelin glycoprotein.";  
 RL J. Cell Biol. 111:2673-2679(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 25-52.  
 RX MEDLINE=88198371; PubMed=3283151;  
 RA Mikol D.D., Stefansson K.;  
 RT "A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein  
 RT in central nervous system myelin and on oligodendrocytes.";  
 RL J. Cell Biol. 106:1273-1279(1988).  
 CC -!- FUNCTION: Cell adhesion molecule contributing to the interactive  
 CC process required for myelination in the central nervous system.  
 CC -!- SUBUNIT: Binds to RTN4R (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -!- TISSUE SPECIFICITY: Oligodendrocytes and myelin of the central  
 CC nervous system.  
 CC -!- PTM: O-Glycosylated in its Ser/Thr-rich repeat domain (Potential).  
 CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.  
 CC -!- CAUTION: Do not confuse oligodendrocyte-myelin glycoprotein (OMG)  
 CC with myelin-oligodendrocyte glycoprotein (MOG).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M63623; AAA59970.1; ALT\_SEQ.  
 DR EMBL; X57436; CAA40684.1; -.  
 DR EMBL; X51694; CAA35991.1; ALT\_SEQ.  
 DR EMBL; BC018050; AAH18050.1; -.  
 DR PIR; A36688; A39613.  
 DR Genew; HGNC:8135; OMG.  
 DR MIM; 164345; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_typ.  
 DR Pfam; PF00560; LRR; 5.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00369; LRR\_TYP; 2.  
 DR SMART; SM00013; LRRNT; 1.  
 KW Glycoprotein; Myelin; Cell adhesion; Repeat; Leucine-rich repeat;  
 KW Signal; GPI-anchor.  
 FT SIGNAL 1 24  
 FT CHAIN 25 440 OLIGODENDROCYTE-MYELIN GLYCOPROTEIN.  
 FT REPEAT 33 54 LRR 1.  
 FT REPEAT 55 77 LRR 2.  
 FT REPEAT 78 100 LRR 3.  
 FT REPEAT 122 145 LRR 4.  
 FT REPEAT 146 168 LRR 5.  
 FT REPEAT 170 189 LRR 6.  
 FT REPEAT 190 213 LRR 7.  
 FT REPEAT 217 228 LRR 8 (INCOMPLETE).  
 FT REPEAT 229 270 SER/THR-RICH.  
 FT REPEAT 271 292 SER/THR-RICH.  
 FT REPEAT 293 335 SER/THR-RICH.

FT	REPEAT	336	377	SER/THR-RICH.
FT	REPEAT	378	416	SER/THR-RICH.
FT	CARBOHYD	45	45	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	61	61	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	103	103	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	189	189	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	192	192	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	389	389	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	425	425	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	440 AA; 49608 MW; 762FD0E8905EA050 CRC64;		

Query Match 59.7%; Score 37; DB 1; Length 440;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPQNFYKLPQ 10  
 || || ::||  
 Db 386 MPNNFSEMPQ 395

# RESULT 15

## BRA6\_CHITH

ID BRA6\_CHITH STANDARD; PRT; 453 AA.  
 AC P24243; Q8MPQ2;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Balbiani ring A 67 kDa protein precursor (p67).  
 GN F6.2.  
 OS Chironomus thummi thummi (Midge).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomoidea;  
 OC Chironomidae; Chironominae; Chironomus.  
 OX NCBI\_TaxID=7155;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND REVISIONS.  
 RC TISSUE=Salivary gland;  
 RA Alieva E., Mayorov V., Elisafenko E., Adkison L., Blinov A.G.;  
 RT "Analysis of structure and expression of F6.2 gene in Chironomus  
 RT thummi and in related species.";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-293 FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=91099682; PubMed=2269433;  
 RA Bogachev S.S., Blinov A.G., Kolesnikov N.N., Scherbik S.V.,  
 RA Taranin A.V., Sebeleva T.E., Baiborodin S.I., Kiknadze I.I.;  
 RT "A tissue-specific puff (Balbiani ring a) in Chironomus thummi may  
 RT contain a gene encoding a 67-kDa protein which exhibits non-tissue-  
 RT specific expression.";  
 RL Gene 96:241-247(1990).  
 CC -!- FUNCTION: Used by the larvae to construct a supramolecular  
 CC structure, the larval tube.

CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: The 67 kDa protein is not tissue-specific.  
 CC Thus, the puff Balbiani ring, which is tissue-specific, contains  
 CC at least one non-tissue-specific gene.  
 CC -!- MISCELLANEOUS: This ORF may only constitute one exon for the 67  
 CC kDa protein.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF521043; AAM77906.1; -.  
 DR PIR; JH0365; JH0365.  
 KW Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 453 BALBIANI RING A 67 kDa PROTEIN.  
 SQ SEQUENCE 453 AA; 52690 MW; D6409D541FC29780 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 453;  
 Best Local Similarity 75.0%; Pred. No. 21;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QNFYKLPQ 10  
 :||:||||  
 Db 65 KNFFKLPQ 72

Search completed: August 24, 2004, 15:43:33  
 Job time : 8.91045 secs